



## APLICAÇÃO DE REDES NEURAIS NO CONTROLE DE TUBERCULOSE BOVINA

Roberto Tadeu de Andrade Filho

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: José Manoel de Seixas

Rio de Janeiro  
Março de 2016

APLICAÇÃO DE REDES NEURAIS NO CONTROLE DE TUBERCULOSE  
BOVINA

Roberto Tadeu de Andrade Filho

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. José Manoel de Seixas, D.Sc.

---

Prof. Afrânio Lineu Kritski, Ph.D.

---

Prof.<sup>a</sup> Aline Gesualdi Manhães, D.Sc.

---

Prof.<sup>a</sup> Marley Maria Bernardes Rebuzzi Vellasco, Ph.D.

---

Eng. Rodrigo de Coura Torres, D.Sc.

---

Prof. Walter Lilenbaum, D.Sc.

RIO DE JANEIRO, RJ – BRASIL

MARÇO DE 2016

Andrade Filho, Roberto Tadeu de

Aplicação de Redes Neurais no Controle de Tuberculose Bovina/Roberto Tadeu de Andrade Filho. – Rio de Janeiro: UFRJ/COPPE, 2016.

XVI, 161 p.: il.; 29,7cm.

Orientador: José Manoel de Seixas

Dissertação (mestrado) – UFRJ/COPPE/Programa de Engenharia Elétrica, 2016.

Referências Bibliográficas: p. 153 – 161.

1. Redes Neurais Artificiais. 2. *MLP*. 3. Classificação Supervisionada. 4. Tuberculose Bovina. I. Seixas, José Manoel de. II. Universidade Federal do Rio de Janeiro, COPPE, Programa de Engenharia Elétrica. III. Título.

*Aos meus pais, maiores  
incentivadores.*

# Agradecimentos

Gostaria de agradecer primeiramente aos meus pais. Por todo o amor, compreensão e carinho que tiveram por mim no período de desenvolvimento deste trabalho.

Ao meu orientador, José Manoel de Seixas por aceitar me orientar. Seus conselhos fizeram com que este trabalho se tornasse realidade.

Ao Professor Flávio Ribeiro, por me ajudar e esclarecer as dúvidas pertinentes quanto a tuberculose bovina e a área de veterinária em geral.

Aos meus amigos, pelo apoio dado.

Por fim, agradeço e dedico este trabalho a Natasha Machado Araujo, amiga, companheira. Seu amor, carinho, paciência, sua ajuda e conselhos foram essenciais para a finalização deste trabalho.

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

## APLICAÇÃO DE REDES NEURAIS NO CONTROLE DE TUBERCULOSE BOVINA

Roberto Tadeu de Andrade Filho

Março/2016

Orientador: José Manoel de Seixas

Programa: Engenharia Elétrica

A tuberculose bovina é uma doença que atinge bovinos e humanos, de evolução lenta, contagiosa, de difícil detecção e sem tratamento para os bovinos. Por esses motivos, ela é considerada preocupante tanto sanitariamente quanto economicamente. No Brasil, existe o programa PNCEBT que tem como objetivo regulamentar as estratégias de controle e erradicação a doença. O programa considera como exame oficial para diagnóstico da Tuberculose Bovina, o exame de Teste Cervical. Sabe-se, porém, que esse exame apresenta limitações, não diagnosticando com precisão o animal doente. Este trabalho analisou dados de três fazendas, de regiões diferentes, e propôs o desenvolvimento de modelos neurais como ferramenta de análise e avaliar o impacto gerado pela adição de novos exames e seu uso em conjunto ao diagnóstico da doença. São também comparados os pontos de corte atuais e oficiais do Teste Cervical recomendados pelo programa nacional aos novos pontos de corte obtidos neste estudo.

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

## NEURAL NETWORKS APLICATION ON THE BOVINE TUBERCULOSIS CONTROL

Roberto Tadeu de Andrade Filho

March/2016

Advisor: José Manoel de Seixas

Department: Electrical Engineering

The bovine tuberculosis is an disease that affects cattle and humans, with slow evolution, being contagious and difficult to detect, as well with no treatment for the cattle. For these reasons, the disease is considered both a health issue and an economic issue. In Brazil, there is the PNCEBT program that aims at regulating the strategies control and eradication of the illness. The program determines as the official exam for the diagnosis of bovine tuberculosis, the cervical tests. It is known, however, that these tests have limitations, not diagnosing accurately the sick animals. This study analyzed data from three different farms from different regions, and proposed the development of neural models as an analytical tool to evaluate the impact generated by the addition of new tests and their use together in diagnosing the disease. The study also compares the current official cutoffs of the cervical tests, recommended by the national program, with the new cutoff points obtained with this study.

# Sumário

<b>Lista de Figuras</b>	<b>xi</b>
<b>Lista de Tabelas</b>	<b>xiv</b>
<b>1 Introdução</b>	<b>1</b>
1.1 Motivação . . . . .	2
1.2 Objetivo . . . . .	3
1.3 Organização do Trabalho . . . . .	3
<b>2 Tuberculose Bovina</b>	<b>5</b>
2.1 Epidemiologia . . . . .	5
2.2 Diagnóstico . . . . .	6
2.2.1 Tuberculinização . . . . .	7
2.2.2 ELISA . . . . .	9
2.2.3 Baciloscopia . . . . .	10
2.2.4 Cultura . . . . .	11
2.2.5 PCR . . . . .	11
2.2.6 Anergia . . . . .	13
2.3 Impacto Econômico . . . . .	13
2.4 Programa de Controle e Erradicação . . . . .	14
<b>3 Técnica de Aprendizado de Máquina</b>	<b>16</b>
3.1 Redes Neurais Artificiais . . . . .	16
3.1.1 Tipos de Aprendizado . . . . .	17
3.1.2 Neurônio Biológico e Neurônio Artificial . . . . .	18
3.1.3 Arquitetura das Redes Neurais Artificiais . . . . .	22
3.1.4 Redes Multicamadas Perceptron . . . . .	26
3.1.5 Algoritmo <i>Backpropagation</i> . . . . .	29
3.1.6 Algoritmo <i>Resilient Backpropagation</i> . . . . .	33
3.1.7 Redes de Operação . . . . .	35
3.2 Métodos Estatísticos . . . . .	36
3.2.1 Acurácia, Especificidade e Sensibilidade . . . . .	36

3.2.2	Curva ROC . . . . .	38
3.2.3	Valor SP . . . . .	42
3.2.4	Validação Cruzada . . . . .	45
<b>4</b>	<b>Método Experimental</b>	<b>48</b>
4.1	Dados . . . . .	48
4.1.1	Fazenda UFF . . . . .	49
4.1.2	Fazenda Macaé . . . . .	51
4.1.3	Fazenda RS . . . . .	52
4.1.4	Preparação dos dados . . . . .	55
4.2	Adição de exames aos testes cervicais . . . . .	59
4.2.1	Os modelos neurais . . . . .	60
4.2.2	Topologia dos modelos . . . . .	61
4.2.3	Treinamento . . . . .	63
4.3	Repetibilidade . . . . .	66
4.3.1	Uso dos dados e escolha dos modelos neurais . . . . .	67
4.3.2	Redes de Operação . . . . .	68
4.4	Otimização dos testes cervicais . . . . .	69
4.4.1	Ajuste da norma oficial . . . . .	70
4.4.2	Análise . . . . .	71
<b>5</b>	<b>Resultados</b>	<b>72</b>
5.1	Adição de exames . . . . .	72
5.1.1	Fazenda UFF . . . . .	72
5.1.2	Fazenda Macaé . . . . .	87
5.1.3	Fazenda RS . . . . .	100
5.1.4	Análise conjunta das fazendas . . . . .	105
5.2	Repetibilidade . . . . .	113
5.2.1	Redes de Operação SEM anérgicos no treinamento . . . . .	114
5.2.2	Redes de Operação COM anérgicos no treinamento . . . . .	119
5.3	Otimização dos testes cervicais . . . . .	124
5.3.1	Fazenda UFF . . . . .	125
5.3.2	Fazenda Macaé . . . . .	131
5.3.3	Fazenda RS . . . . .	136
5.3.4	Fazenda All . . . . .	141
<b>6</b>	<b>Conclusões</b>	<b>148</b>
6.1	Adição de exames e Repetibilidade . . . . .	148
6.2	Testes Cervicais . . . . .	150
6.3	Trabalhos Futuros . . . . .	152



# Lista de Figuras

3.1	Estrutura do Neurônio Biológico [1] . . . . .	18
3.2	Modelo de Neurônio Artificial de McCulloch e Pitts . . . . .	20
3.3	Formato das Funções de Ativação consideradas . . . . .	21
3.4	Modelo de Neurônio Artificial . . . . .	22
3.5	Redes Neurais e seus tipos de conexão . . . . .	23
3.8	Formato de uma rede MLP . . . . .	27
3.9	Exemplo de Erros em treinamento de rede MLP . . . . .	29
3.10	Exemplo de convergência do Vetor Gradiente . . . . .	30
3.11	Resultados possíveis para as condições Negativa e Positiva . . . . .	36
3.12	Exemplos de curva ROC . . . . .	39
3.13	Exemplos de pontos no espaço de uma curva ROC . . . . .	40
3.14	Exemplos de medidas <b>AUC</b> em curvas ROC . . . . .	41
3.15	Exemplo de Integração pela Regra do Trapézio . . . . .	42
3.16	Variação do Valor SP baseado nos valores da Sensibilidade e da Especificidade . . . . .	43
3.17	Distribuição e Resultados do cenário hipotético com análise por ROC e SP [2] . . . . .	44
3.18	Método <i>Holdout</i> de Validação Cruzada . . . . .	45
3.19	Método <i>K-fold</i> de Validação Cruzada . . . . .	46
4.1	Exemplo de Topologia de modelo neural desenvolvido . . . . .	62
4.2	Exemplo de evolução do erro de validação de modelo neural . . . . .	64
5.1	Distribuição do SP para os modelos da fazenda UFF <b>SEM</b> animais anérgicos no treinamento . . . . .	73
5.2	Gráficos de curva ROC para a fazenda UFF, SEM anérgicos no treinamento . . . . .	74
5.3	Média e Desvio Padrão das medidas de performance para a fazenda UFF, <b>SEM</b> anérgicos no treinamento . . . . .	76
5.4	Sensibilidade dos animais anérgicos para os modelos da fazenda UFF <b>SEM</b> animais anérgicos no treinamento . . . . .	78

5.5	Distribuição do SP para os modelos da fazenda UFF <b>COM</b> animais anérgicos no treinamento . . . . .	79
5.6	Gráficos de curva ROC para a fazenda UFF, <b>COM</b> anérgicos no treinamento . . . . .	80
5.7	Média e Desvio Padrão das medidas de performance para a fazenda UFF, <b>COM</b> anérgicos no treinamento . . . . .	82
5.8	Sensibilidade dos animais anérgicos para os modelos da fazenda UFF <b>COM</b> animais anérgicos no treinamento . . . . .	84
5.9	Distribuição do SP para os modelos da fazenda Macaé <b>SEM</b> animais anérgicos no treinamento . . . . .	88
5.10	Gráficos de Curva ROC para a fazenda Macaé, <b>SEM</b> Anérgicos no Treinamento . . . . .	89
5.11	Média e Desvio Padrão das medidas de performance para a fazenda Macaé, <b>SEM</b> anérgicos no treinamento . . . . .	91
5.12	Sensibilidade dos animais anérgicos para os modelos da fazenda Macaé <b>SEM</b> animais anérgicos no treinamento . . . . .	92
5.13	Distribuição do SP para os modelos da fazenda Macaé <b>COM</b> animais anérgicos no treinamento . . . . .	93
5.14	Gráficos de Curva ROC para a fazenda Macaé, <b>COM</b> anérgicos no treinamento . . . . .	94
5.15	Média e Desvio Padrão das medidas de performance para a fazenda Macaé, <b>COM</b> anérgicos no treinamento . . . . .	96
5.16	Sensibilidade dos animais anérgicos para os modelos da fazenda Macaé <b>COM</b> animais anérgicos no treinamento . . . . .	97
5.17	Distribuição do SP para os modelos da fazenda RS <b>SEM</b> animais anérgicos no treinamento . . . . .	100
5.18	Gráficos de Curva ROC para a fazenda RS, <b>SEM</b> anérgicos no treinamento . . . . .	101
5.19	Média e Desvio Padrão das medidas de performance para a fazenda RS, <b>SEM</b> anérgicos no treinamento . . . . .	103
5.20	Média e Desvio Padrão das medidas de performance para todas as fazendas, <b>SEM</b> anérgicos no treinamento . . . . .	106
5.21	Distribuição do SP para os modelos da fazenda RS <b>SEM</b> animais anérgicos no treinamento . . . . .	108
5.22	Média e Desvio Padrão das medidas de performance para todas as fazendas, <b>COM</b> anérgicos no treinamento . . . . .	110
5.23	Distribuição do SP para os modelos da fazenda RS <b>SEM</b> animais anérgicos no treinamento . . . . .	112

5.24 Saídas das Redes de Operação para os animais de RS, <b>SEM</b> anérgicos no treinamento . . . . .	116
5.25 Dados da fazenda RS simulados nos modelos neurais de UFF e Macaé treinados <b>SEM</b> anérgicos . . . . .	118
5.26 Saídas das Redes de Operação para os animais de RS, <b>COM</b> anérgicos no treinamento . . . . .	121
5.27 Dados da fazenda RS simulados nos modelos neurais de UFF e Macaé treinados <b>COM</b> anérgicos . . . . .	123
5.28 Variação do SP pelo ponto de corte da Fazenda UFF para os exames TCS e TCC . . . . .	125
5.29 Curvas ROC da Fazenda UFF para os exames TCS e TCC . . . . .	127
5.30 Testes Cervicais de UFF para animais não anérgicos . . . . .	129
5.31 Testes Cervicais de UFF para animais anérgicos . . . . .	130
5.32 Variação do SP pelo ponto de corte da Fazenda Macaé para os exames TCS e TCC . . . . .	132
5.33 Curvas ROC da Fazenda Macaé para os exames TCS e TCC . . . . .	133
5.34 Testes Cervicais de Macaé para animais não anérgicos . . . . .	135
5.35 Testes Cervicais de Macaé para animais anérgicos . . . . .	136
5.36 Variação do SP pelo ponto de corte da Fazenda RS para os exames TCS e TCC . . . . .	137
5.37 Curvas ROC da Fazenda RS para os exames TCS e TCC . . . . .	138
5.38 Testes Cervicais de RS para animais não anérgicos . . . . .	140
5.39 Variação do SP pelo ponto de corte para todas as Fazendas ( <i>All</i> ) para os exames TCS e TCC . . . . .	142
5.40 Curvas ROC da Fazenda <i>All</i> para os exames TCS e TCC . . . . .	143
5.41 Testes Cervicais ma Fazenda <i>All</i> para animais não anérgicos . . . . .	145
5.42 Testes Cervicais ma Fazenda <i>All</i> para animais anérgicos . . . . .	146

# Lista de Tabelas

2.1	Interpretação do Teste Cervical Simples em bovinos [3] . . . . .	8
2.2	Interpretação do Teste Cervical Comparativo em bovinos [3] . . . . .	8
3.1	Possíveis resultados entre a observação e a real condição . . . . .	37
4.1	Relação de valores e significados dos campos dos conjuntos de dados .	49
4.2	Quantidade de Animais TB+ e TB− para a fazenda Fazenda UFF . .	49
4.3	Relação de exames usados, tipo do dado e seus possíveis valores - UFF	50
4.4	Quantidade de Animais TB+ e TB− para a fazenda Fazenda Macaé .	51
4.5	Relação de exames usados, tipo do dado e seus possíveis valores - Macaé	52
4.6	Quantidade de Animais TB+ e TB− para Fazenda RS . . . . .	53
4.7	Relação de exames usados, tipo do dado e seus valores - RS . . . . .	54
4.8	Relação de animais mantidos e animais removidos por fazenda . . . .	56
4.9	Representação do diagnóstico na Rede Neural . . . . .	57
4.10	Interpretação dos valores normalizados de ELISA . . . . .	58
4.11	Relação dos modelos neurais por fazenda . . . . .	61
4.12	Proporção dos dados para cada grupo . . . . .	65
4.13	Relação dos modelos neurais usados para a propagação de RS . . . .	67
5.1	Performance das Redes neurais de UFF para animais não anérgicos. Treinamento <b>SEM</b> anérgicos . . . . .	77
5.2	Performance das Redes neurais de UFF para animais anérgicos. Trei- namento <b>SEM</b> anérgicos . . . . .	78
5.3	Performance das Redes neurais de UFF para animais não anérgicos. Treinamento <b>COM</b> anérgicos . . . . .	83
5.4	Performance das Redes neurais de UFF para animais anérgicos. Trei- namento <b>COM</b> anérgicos . . . . .	84
5.5	Performance dos exames isolados de UFF para animais não anérgicos	85
5.6	Performance de identificação de animais não anérgicos testados com o exame de leite . . . . .	86
5.7	Performance dos exames isolados de UFF para animais anérgicos . . .	86

5.8	Performance das redes neurais de Macaé para animais não anérgicos. Treinamento <b>SEM</b> anérgicos . . . . .	90
5.9	Performance das redes neurais de Macaé para animais anérgicos. Trei- namento <b>SEM</b> anérgicos. . . . .	93
5.10	Performance das redes neurais de Macaé para animais não anérgicos. Treinamento <b>COM</b> anérgicos . . . . .	97
5.11	Performance das redes neurais de Macaé para animais anérgicos. Trei- namento <b>COM</b> anérgicos . . . . .	98
5.12	Performance dos exames isolados de Macaé para animais não anérgicos	98
5.13	Performance dos exames isolados de Macaé para animais anérgicos . .	99
5.14	Performance das redes neurais de RS para animais não anérgicos. Treinamento <b>SEM</b> anérgicos . . . . .	104
5.15	Performance dos exames isolados de RS para animais não anérgicos .	104
5.16	Comparação dos modelos neurais entre todas as fazendas, somente para animais não anérgicos. Treinamento dos modelos <b>SEM</b> anérgicos	108
5.17	Comparação dos modelos entre as fazendas UFF e Macaé, somente animais anérgicos. Treinamento dos modelos <b>SEM</b> anérgicos . . . . .	109
5.18	Comparação dos modelos neurais entre todas as fazendas, somente para animais não anérgicos. Treinamento dos modelos <b>COM</b> anérgicos	112
5.19	Comparação dos modelos entre as fazendas UFF e Macaé, somente animais anérgicos. Treinamento dos modelos <b>COM</b> anérgicos . . . . .	113
5.20	Redes de Operação para os grupos de treinamento e teste, <b>SEM</b> anérgicos no treinamento . . . . .	114
5.21	Dados de RS propagados nas redes de operação treinadas <b>SEM</b> Anérgicos . . . . .	117
5.22	Redes de Operação para os grupos de treinamento e teste, <b>COM</b> anérgicos no treinamento . . . . .	119
5.23	Dados de RS propagados nas redes de operação treinadas <b>COM</b> anérgicos . . . . .	122
5.24	Valores de Faixa dinâmica, Pontos de Corte e SP para TCS e TCC da fazenda UFF . . . . .	126
5.25	Eficiência das normas para exames cervicais para a fazenda UFF - Animais Não Anérgicos . . . . .	128
5.26	Eficiência das normas para cervicais para a fazenda UFF - Animais Anérgicos . . . . .	129
5.27	Valores de Faixa dinâmica, Pontos de Corte e SP para TCS e TCC da fazenda Macaé . . . . .	131
5.28	Eficiência das normas para exames cervicais para a fazenda Macaé - Animais Não Anérgicos . . . . .	134

5.29	Eficiência das normas para exames cervicais para a fazenda Macaé - animais anérgicos . . . . .	135
5.30	Valores de Faixa dinâmica, Pontos de Corte e SP para TCS e TCC da fazenda RS . . . . .	137
5.31	Eficiência das normas para exames cervicais para a fazenda RS - animais não anérgicos . . . . .	139
5.32	Quantidades de Animais TB e Não TB na fazenda <i>All</i> . . . . .	141
5.33	Valores de Faixa dinâmica, Pontos de Corte e SP para TCS e TCC da fazenda <i>All</i> . . . . .	142
5.34	Eficiência das normas para exames cervicais para a fazenda <i>All</i> - animais não anérgicos . . . . .	144
5.35	Eficiência das normas para exames cervicais para a fazenda <i>All</i> - animais anérgicos . . . . .	146

# Capítulo 1

## Introdução

Existem diversas doenças que atingem animais e são transmitidas para os seres humanos. Das doenças que acometem bovinos destaca-se a tuberculose bovina, que é uma zoonose causada pela micobactéria *M. bovis* e tem como característica o aparecimento de lesões, principalmente nos pulmões, mas também em outras partes do corpo do animal [4]. Ela é uma doença cujo primeiro hospedeiro é o bovino, e que é transmitida principalmente pela via aérea. Tanto animais que convivem no mesmo rebanho podem ser afetados pela doença assim como profissionais que trabalham e têm contato intenso com os mesmos. A doença também pode ser encontrada em outras espécies de animais [5].

A tuberculose bovina é uma doença crônica e se manifesta lentamente, podendo demorar meses ou anos desde a contaminação até que apareçam os primeiros sintomas. Devido a esse fato, em geral, quando se percebe que o animal está enfermo, a doença já está muito avançada e ele precisa ser sacrificado, além da grande possibilidade de contágio de outros animais.

Uma vez que seu diagnóstico clínico é demorado, é importante que haja outras formas de identificar se um animal está contaminado em um estágio menos avançado da doença, para evitar a contaminação dos produtos de origem bovina e também dos demais animais. São utilizados exames no animal afim de obter um diagnóstico precoce. O exame recomendado pela legislação brasileira é o exame de Teste Cervical [3], porém ele não é preciso e apresenta limitações no diagnóstico.

Em virtude disso, vem sendo pesquisados e utilizados outros exames, que combinados podem indicar a presença da doença mais rapidamente. Esses exames são sorológicos, bacteriológicos e moleculares, por exemplo.

Devido à dificuldade do diagnóstico, inexistência de tratamento, além dos danos causados pela doença, a tuberculose bovina se torna não só um sério problema de saúde, como também um sério problema econômico. O animal infectado produz menos leite, precisa ser sacrificado e pode infectar um rebanho inteiro, gerando altos custos para o produtor, e também gerando uma perda do ganho obtido com a venda

de carne e leite. Além disso, uma alta incidência da doença no rebanho brasileiro pode comprometer a exportação de carne também para outros países, gerada pela má reputação da qualidade dos produtos de origem bovina.

Com o objetivo de regulamentar as estratégias de erradicação e diminuição de incidência da doença, foi criado no ano de 2001 pelo governo federal, o Programa Nacional de Controle e Erradicação da Brucelose e Tuberculose (PNCEBT)[3]. O programa divulga as melhores práticas para os exames preventivos no rebanho, além de regulamentar como lidar com a doença uma vez contraída. O programa detalha medidas para proteger tanto os demais animais quanto os profissionais envolvidos, como também visa assegurar a qualidade dos produtos gerados a partir deles.

Por outro lado, o avanço de novas tecnologias apresenta ferramentas capazes de auxiliar no diagnóstico da doença. As Redes Neurais Artificiais [6] [7] são inspiradas no comportamento cognitivo humano e são capazes de aprender a partir de exemplos.

As Redes Neurais Artificiais são aplicadas também na solução de outros problemas médicos [8] como no diagnóstico da Tuberculose pleural [9], em doenças do sistema nervoso central como a Doença Meningocócica [10] e em doenças ginecológicas [11]. Suas aplicações se estendem além da área médica, sendo também aplicadas em diversos problemas, como por exemplo, o reconhecimento de faces humanas [12], a detecção de partículas [13] e a detecção de fraudes com cartões de crédito [14].

## 1.1 Motivação

A tuberculose bovina motiva o estudo por uma série de aspectos, como o problema sanitário causado por ela que implica na contaminação tanto de bovinos como de humanos, necessitando ainda de políticas de prevenção mais fortes. Além disso, o impacto econômico gerado pela doença é preocupante, uma vez que a contaminação de animais afeta diretamente os custos de produção e o potencial de exportação de produtos bovinos.

Outro ponto importante refere-se às limitações nos métodos oficiais do PNCEBT, uma vez que os exames oficiais são muitas vezes falhos e insuficientes para detecção da doença, principalmente em estágios iniciais. Contudo, estudos mostram que uma abordagem multidisciplinar de diagnóstico apresenta resultados positivos e ganhos na detecção de animais infectados [15] [16].

Por fim, as técnicas de Aprendizado de Máquina (*Machine Learning*) e seu potencial de resolução de problemas complexos [17], motivam este estudo em buscar aplicar tais técnicas ao problema da tuberculose bovina.

## 1.2 Objetivo

Dado os aspectos expostos na seção anterior, o objetivo deste trabalho é promover um estudo sobre as possibilidades que possam aumentar a eficácia dos métodos de diagnóstico da tuberculose bovina, hoje vigente no país.

O presente estudo aborda o impacto causado pela incorporação de novos exames aos já preconizados pelo PNCEBT e discute-se se existem benefícios para tal inserção de exames. As Redes Neurais Artificiais, por se mostrarem aplicáveis em problemas da área médica, foram escolhidas como ferramenta para a avaliação da inserção de novos testes ao diagnóstico da tuberculose bovina. O objetivo deste trabalho é verificar, dentre os exames disponíveis e analisados, quais são os candidatos a serem incluídos no programa e definir uma ordem para sua implantação.

Este estudo propõe ainda uma norma alternativa para os Testes Cervicais, diferente da norma oficial definida pelo governo. O objetivo é verificar se os Testes Cervicais podem apresentar uma maior eficácia, representada por uma maior taxa na identificação da doença, quando comparadas com as normas definidas pelo PNCEBT.

## 1.3 Organização do Trabalho

Este trabalho está dividido em seis capítulos. O primeiro capítulo é a presente introdução deste trabalho.

O segundo capítulo trás a revisão bibliográfica da doença. São apresentados os conceitos acerca da doença, como o seu agente causador, sua forma de contágio e os sintomas típicos que ela apresenta. O capítulo ainda discute sobre os diferentes métodos de diagnóstico e exames confirmatórios, além dos impactos econômicos causado pela tuberculose bovina.

O terceiro capítulo aborda o campo de *Machine Learning*, com ênfase na técnica usada neste trabalho, Redes Neurais Artificiais. É apresentado um breve histórico e sua evolução, seus conceitos básicos e características mais importantes. Por fim, o capítulo expõe um conjunto de técnicas estatísticas usadas neste trabalho, para análise dos modelos neurais e dos dados.

No quarto capítulo são descritos os dados utilizados no experimento, assim como o método adotado, a arquitetura das redes neurais desenvolvidas, assim como um detalhamento da condução do experimento e a aplicação das técnicas descritas anteriormente.

No quinto capítulo são apresentados os resultados gerados a partir da aplicação das técnicas, assim como suas comparações e análise dos dados.

O sexto e último capítulo apresenta as conclusões obtidas pela aplicação do pro-

posto experimento e dos resultados obtidos. O capítulo traz ainda uma discussão quanto ao incremento de exames aos que já são hoje usados no controle da tuberculose bovina. Por fim, apresenta-se possibilidades de trabalhos futuros e perspectivas para a pesquisa.

# Capítulo 2

## Tuberculose Bovina

A tuberculose bovina é uma doença infectocontagiosa de caráter crônico, que apresenta como agente etiológico o agente *Mycobacterium bovis* [3]. Ela pode atingir bovinos, caprinos, ovinos, suínos, animais silvestres e também humanos [18]. Além de representar um problema de saúde pública, ela pode representar também um sério problema econômico, uma vez que influencia na comercialização dos produtos de fonte bovina [19].

Neste capítulo serão descritas as principais características da tuberculose bovina e seus principais exames diagnósticos.

### 2.1 Epidemiologia

Considerada um grave problema de saúde pública, a Tuberculose Bovina é uma doença de evolução lenta, causada pela micobactéria *M. bovis*, e caracteriza-se pelo desenvolvimento de lesões nodulares que podem ser localizadas em qualquer órgão do animal [20]. Devido à natureza de sua criação livre e do contato direto com os demais animais, os rebanhos leiteiros possuem um risco maior de contaminação, porém em algumas regiões a doença é observada também em rebanhos de corte, principalmente em propriedades com maior concentração de animais e regiões alagadas [21].

Os bovinos podem se contaminar de forma direta ou indireta [19]. A transmissão pode ocorrer por meio de gotículas de *Pflugge* em suspensão, inalação de pó com o bacilo e pelas vias respiratórias, sendo a última a forma mais comum de infecção, correspondendo a 90% dos casos de animais adultos [22] [23]. Outras formas de contágio se dão através de leite infectado ingerido por bezerros, transmissão pela placenta e intra uterina pelo coito, porém as duas últimas são mais raras [19].

É importante observar que o bovino infectado pode transmitir o agente *M. bovis* antes mesmo que apareça qualquer tipo de lesão nos tecidos ou sintomas clínicos [23]. O agente causador da doença pode ser encontrado em diversos órgãos e partes

do corpo do animal [24], sendo porém, mais frequentemente encontrado na região dos pulmões.

Uma grande preocupação com relação à doença é a contaminação do homem ao ingerir um alimento de origem bovina, como o leite ou carne. Fatores como insolação, baixa umidade e altas temperaturas restringem a sobrevivência do agente *M. bovis* [3]. O processo de pasteurização do leite, minimiza essa forma de transmissão da tuberculose bovina. Para a carne bovina, o risco de transmissão se dá pela ingestão de carne contaminada e mal cozida. Logo, inspeções realizadas pelos frigoríficos se destacam como o método que assegura a qualidade da carne [21].

Outra forma de contágio da tuberculose bovina é pela inalação de aerossóis virulentos de *M. bovis*, expelidos pelos animais tuberculosos. Tal forma de contágio é comum entre os profissionais que atuam ou que tenham contato direto com o rebanho ou com produtos provenientes deste, como criadores, tratadores e veterinários [23] [19].

O tratamento da tuberculose bovina em humanos é semelhante ao tratamento da tuberculose causada pelo agente *M. tuberculosis*, já que clinicamente as doenças não são distinguíveis. Já nos bovinos, devido à demora na manifestação da doença e no possível diagnóstico, o tratamento é proibido [22]. O PNCEBT baseou parte de sua estratégia de controle e erradicação em países que obtiveram sucesso por meio de seus próprios programas. Em tais programas, não foram conduzidos tratamentos para os animais e por este motivo, o PNCEBT adota as mesmas práticas [3].

A legislação indica que os animais suspeitos ou já confirmados com a doença sejam mantidos afastados do rebanho por no máximo trinta dias, e ao fim desse prazo sejam sacrificados. O sacrifício desse animal deve seguir uma série de diretrizes, sendo realizado em um local oficial ou sob fiscalização, e deve-se tomar cuidado com a contaminação das pessoas envolvidas no transporte e manejo, assim como o local de enterro [3]. Somente funcionários autorizados e com equipamentos de proteção adequados podem conduzir os animais por tais etapas seguintes à identificação da doença.

## 2.2 Diagnóstico

Os estágios iniciais da tuberculose bovina usualmente não apresentam sintomas detectáveis, e podem demorar alguns meses até que surja o primeiro sintoma. Infecções podem também não se manifestar por vários anos, sendo ativadas apenas quando o animal fica estressado ou envelhece. O primeiro sintoma notado pode ser uma tosse que se torna crônica depois que a bactéria se instalou completamente no animal [25] [20].

Embora a infecção seja mais comum nos pulmões, ela também pode afetar o

cérebro, o fígado, os rins ou a medula espinhal [25]. Com o avanço da doença, o animal pode apresentar redução de peso devido à falta de apetite, e pode se tornar mais fraco e debilitado. Quando os animais chegam ao estágio terminal, eles podem apresentar dificuldades respiratórias graves [25] [20]. Essa infecção pode levar a uma redução de até 25% na produção de leite e carne [19].

Já no homem, a doença pode atingir os pulmões, causando febre, tosse, dores no peito e nas costas, emagrecimento ou ainda atingir rins, articulações, fígado entre outros [20].

Desta forma, uma vez que a doença se manifesta muito lentamente, os sinais para observação só se tornam evidentes quando ela já está muito avançada e o animal pode ter infectado também outros animais e humanos. Logo, é importante que existam formas de detectar essa doença mais prematuramente. Existem então exames que, se feitos regularmente, conseguem detectar as infecções ainda no estágio inicial [25] [20]. Alguns deles serão vistos a seguir.

### 2.2.1 Tuberculinização

O teste mais utilizado enquanto o animal ainda está vivo é a tuberculinização. Ela pode ser basicamente de três formas: caudal, cervical simples ou cervical comparada. O exame é o recomendado para diagnosticar a tuberculose bovina *in vivo*, apesar de possuir limitações [26] [27]. A tuberculinização deve ser realizada por um médico veterinário capacitado e credenciado pelo Ministério da Agricultura. Atualmente, o órgão divulga que aproximadamente 5600 médicos veterinários estão credenciados pelo PNCEBT, estando estes divididos por estado.

A prova da tuberculina possui um princípio simples, onde uma vez exposto ao agente, o animal desenvolve resposta imune específica. A tuberculina é injetada intradermicamente no animal e ele reage gerando um aumento de volume no local da injeção. Dependendo do tamanho e cor dessa região, pode se caracterizar uma reação positiva [26]. O Derivado Proteico Purificado (PPD), obtido de cultura da amostra AN5 de *M. bovis*, é a tuberculina mais utilizada.

No Teste Caudal, inicialmente é realizada a medição da espessura da pele da prega da cauda do animal com o cutímetro. Em seguida, é injetado 0,1 ml de tuberculina na região e, após 72 horas, o veterinário realiza a medição analisando se houve alguma alteração no local ou não.

O Teste Cervical Simples (TCS) é feito de maneira semelhante ao teste caudal, injetando o PPD bovino e fazendo a comparação após 72 horas. Porém, nesse exame a região do animal onde ocorre a injeção é a tábua do pescoço [26]. A avaliação do TCS é feita considerando o aumento da espessura da dobra da pele onde se aplicou o PPD bovino. A Tabela 2.1 trás a relação das reações e seus diagnósticos, segundo

a definição estabelecida pelo governo. Uma reação negativa é considerada para um aumento da espessura da pele inferior a 2,0 mm. Já uma reação inconclusiva ocorre quando o aumento é superior ou igual a 2,0 mm e inferior a 4,0 mm, sem exsudação ou necrose; e uma reação positiva pode ser dada a partir de 4,0 mm de aumento dependendo dos outros valores apresentados na Tabela 2.1 [3].

Tabela 2.1: Interpretação do Teste Cervical Simples em bovinos [3]

$\Delta Bov(mm)$	Sensibilidade	Consistência	Alterações	Interpretação
$0 \geq \Delta Bov < 2$	-	-	-	Negativo
$2 \geq \Delta Bov < 4$	pouca dor	endurecida	delimitada	Inconclusivo
$2 \geq \Delta Bov < 4$	muita dor	macia	exsudato, necrose	Positivo
$\Delta Bov \geq 4$	-	-	-	Positivo

O problema desse teste é que apesar de simples e barato, ele tende a gerar falsos positivos, em função das semelhanças entre os agentes *M. bovis* e *M. avium* (causador da tuberculose aviária), fazendo com que o animal reaja mesmo que não haja presença do agente bovino [26].

Visando diagnosticar a doença com mais eficácia, é feito, então, o Teste Cervical Comparativo (TCC). Para a realização desse teste, o veterinário deve injetar simultaneamente, mas em locais separados da pele, a PPD bovina e a PPD aviária, e após 72 horas avaliar as reações. O volume das duas regiões é então analisado e a diferença de resultado é comparada conforme a Tabela 2.2 [26] [3].

Tabela 2.2: Interpretação do Teste Cervical Comparativo em bovinos [3]

	$\Delta Bov - \Delta Av(mm)$	Interpretação
$\Delta Bov < 2$	-	Negativo
$\Delta Bov < \Delta Av$	$(\Delta Bov - \Delta Av) < 0$	Negativo
$\Delta Bov \geq \Delta Av$	$0 \geq (\Delta Bov - \Delta Av) < 2$	Negativo
$\Delta Bov > \Delta Av$	$2 \geq (\Delta Bov - \Delta Av) < 4$	Inconclusivo
$\Delta Bov > \Delta Av$	$(\Delta Bov - \Delta Av) \geq 4$	Positivo

Recomenda-se a realização da prova comparativa em casos de rebanhos com muitos resultados inconclusivos no teste cervical simples [3]. É importante que animais suspeitos ou positivos no teste caudal ou cervical simples sejam mantidos sob quarentena até que o cervical comparativo seja realizado [26].

Alguns fatores que podem influenciar no resultado do teste são o armazenamento errado da tuberculina, a sua aplicação imprópria, o animal estar em período de gestação ou ter parido recentemente, e a aplicação de outras substâncias. Além disso, o intervalo entre os testes deve ser de pelo menos 60 dias e no máximo 90 dias [3], para que a capacidade do animal de responder ao teste seja recuperada.

Falsos negativos também podem ser um problema na performance dos testes de tuberculina. Estes são mais comumente observados em estágios iniciais da infecção ou em condições de comprometimento do sistema imune [26].

## 2.2.2 ELISA

O exame *Ensaio de Imunoadsorção Enzimática* (ELISA) é um teste sorológico e imunoenzimático utilizado no diagnóstico de doenças que induzem a produção de anticorpos [28]. O exame permite a detecção de anticorpos específicos presentes na amostra observada no teste, sendo aplicado comumente por exemplo, na identificação de AIDS [29] e de Hepatite B [30] em humanos.

O teste ELISA é feito usando um antígeno recombinante, que reage com os anticorpos presentes no soro do bovino com tuberculose. Essa reação gera o acoplamento da enzima ao anticorpo, que ao adicionar substrato resulta em mudança de cor para as amostras positivas. Essas amostras apresentam uma única cor ao reagirem ao substrato e a intensidade dessa coloração é medida por um aparelho, que interpreta se o resultado é positivo ou negativo [31].

O teste possui aplicação para a tuberculose bovina, sendo realizado nos animais ainda vivos através de uma amostra de sangue, e vem se mostrando um bom exame complementar para a identificação de rebanhos infectados com a doença [31].

Os antígenos mais comuns em testes com ELISA são os antígenos MPB70, MPB83 e ESAT-6. Eles têm sido amplamente pesquisados e testados, porém sua Sensibilidade e Especificidade são baixas. Os anticorpos anti-MPB83 podem ser detectados na fase inicial da infecção, cerca de quatro semanas pós infecção. Já os anticorpos para MPB70 são detectados nos estágios mais avançados da doença, de 18 a 22 meses após a infecção. Os anticorpos anti-ESAT-6 podem ser detectados 12 semanas após a infecção [31].

Foi desenvolvido, então, por pesquisadores da Empresa Brasileira de Pesquisa Agropecuária (EMBRAPA) o exame ELISA Quimérico, que é composto por fragmentos hidrofílicos dos genes desses três antígenos. Em experimentos utilizando amostras com cada antígeno individualmente e a quimera, foi observado que esta apresentou um resultado consideravelmente melhor do que dos antígenos separados, possuindo Sensibilidade e Especificidade superiores [31].

Assim, espera-se que quimeras contendo importantes epítomos das principais

proteínas de *M. bovis* sejam capazes de detectar anticorpos nos bovinos em diferentes fases da infecção, aumentando a cobertura diagnóstica.

O teste ELISA apresenta vantagens como a extração sanguínea do animal em uma única vez, podendo os resultados serem obtidos em 24 horas, além de poder diagnosticar animais em diferentes estágios de infecção, acelerando o combate na propriedade com tuberculose.

### 2.2.3 Baciloscopia

Em 1882, Paul Ehrlich tornou oficial a propriedade do bacilo da tuberculose de absorver coloração e mantê-la, ao deixar bacilos corando em violeta de metila [32]. Essa característica, chamada de ácido resistência, faz com que um bacilo colorido pela substância não perca a coloração após lavagem com ácido nítrico.

Esse fato faz com que as micobactérias sejam classificadas como Bacilos Álcool-Ácido Resistentes, que a prática clínica e laboratorial identifica apenas como BAAR [33]. A baciloscopia, então, é uma técnica onde é observada a presença de bacilos em uma amostra extraída do animal.

Com o desenvolvimento da infecção, pode-se começar a perceber a presença da micobactéria em diversas partes do animal, como em amostras teciduais e com a produção de escarro, devido a um aumento da inflamação e a necrose do tecido pulmonar.

O exame é realizado da seguinte forma: é extraída uma amostra do animal e esta é preparada em esfregaços. O esfregaço é colorido com a substância corante e, utilizando um microscópio, pesquisa-se a presença de BAAR [33]. Uma vez encontrado conclui-se que o animal está infectado. Com isso, a baciloscopia é o método prioritário de triagem, diagnóstico e controle durante o tratamento da tuberculose [34].

Alguns métodos de coloração são conhecidos, porém o método *Ziehl-Neelsen* é um dos mais utilizados, principalmente em localidades com poucos recursos [33]. Ele consiste em corar o bacilo com fucsina e ácido sulfúrico, o que gera a reação de bacilos vermelhos brilhante sobre um fundo azul [18].

A baciloscopia é um exame simples, seguro, rápido e econômico, porém necessita de 5.000 a 10.000 bacilos/mL de amostra para poder ser identificado com precisão, podendo ocorrer assim muitos falsos negativos. Além disso, ele identifica que existe um bacilo do gênero *Mycobacterium*, mas não especifica se trata-se de *M. bovis* ou *M. avium* [34] [4]. O teste de Baciloscopia apresenta uma concordância próxima de 50% com resultados positivos aos testes cervicais [35], ou seja, em uma amostra de animais reagentes à tuberculinização, aproximadamente metade apresenta resultado confirmado pela baciloscopia.

## 2.2.4 Cultura

Outro exame muito utilizado para diagnóstico da tuberculose bovina é o exame de cultura. O exame de cultura consiste em realizar a inoculação dos espécimes em meios de cultura que contenham, além de ingredientes adequados ao crescimento ótimo das micobactérias, um corante capaz de inibir a microbiota de contaminação [18]. Resumidamente, o processo para a realização desse exame é coletar a amostra do animal, descontaminá-la, manter em um ambiente adequado para que ele possa se desenvolver e observar então a presença do agente *M. bovis*.

Os principais fatores que influenciam no sucesso do exame são o meio de cultura utilizado, o processo de descontaminação e as condições de incubação [36].

As micobactérias têm a temperatura ideal de crescimento de 37°C em pH de 6,5. O meio de cultura ideal para a detecção de micobactérias deve oferecer condições para seu crescimento rápido, permitir diferenciação entre micobactérias patogênicas e não patogênicas, inibir o crescimento de contaminantes e ser preparado a partir de ingredientes prontamente disponíveis para sua reprodutibilidade. Porém, a maioria dos meios utilizados não satisfazem à grande parte destas características, pois possuem em sua composição materiais orgânicos complexos [37].

A cultura possui Sensibilidade maior que a baciloscopia, uma vez que pode multiplicar o agente a partir de 10 bacilos/mL. Desta forma, a cultura assume uma grande importância para a detecção de BAAR em espécimes extrapulmonares, e dos suspeitos de tuberculose pulmonar persistentemente negativos à baciloscopia. Adicionalmente, o cultivo permite posterior identificação da micobactéria isolada, assim como a realização de testes de sensibilidade a antibióticos [36] [38].

As principais desvantagens desse exame são o longo tempo necessário para o surgimento de colônias macroscopicamente visíveis e a necessidade de possuir organismos viáveis para sua realização, havendo a possibilidade de inviabilizar a análise dada alguma falha na fase de descontaminação e cultura [39].

## 2.2.5 PCR

Como visto anteriormente, os exames bacteriológicos e de cultura podem apresentar um diagnóstico definitivo da doença, ou seja, uma vez que tais exames respondam positivamente para a presença de *M. bovis*, confirma-se que o animal está infectado. Apesar dessa característica, esses exames apresentam uma série de restrições. As técnicas bacteriológicas, por exemplo, apresentam baixa Sensibilidade, necessitando de uma quantidade muito grande de bacilos para que o agente seja identificado, e a técnica de cultura demora um longo período até seu crescimento ideal. Além disso, a descontaminação pode matar alguns bacilos e reduzir ainda mais sua quantidade [40].

Outra questão é que esses testes identificam apenas o gênero *Mycobacterium*. A identificação entre *M. bovis* e *M. avium* somente é possível por meio de testes bioquímicos e de sensibilidade à drogas, os quais demandam um longo tempo e podem ser falhos [41].

Foi desenvolvida, então, a técnica chamada Reação em Cadeia da Polimerase (*Polymerase Chain Reaction*), com sua sigla definida como PCR e derivada do inglês. A PCR é uma técnica molecular que permite que o diagnóstico seja obtido de forma mais precisa [40] e permite a amplificação *in vitro* de um trecho de DNA milhões de vezes em poucas horas. Ela reproduz o processo natural de replicação do DNA e mostra a ausência ou presença do material genético procurado. Com ela é possível determinar o trecho de DNA que se quer encontrar, e ela funciona mesmo se a quantidade de bacilos for muito pequena [42].

É necessário que a fita a ser duplicada possua alguns nucleotídeos complementares pareados no início para começar a duplicação. A separação das fitas ocorre pela desnaturação térmica do DNA. Para isto, termocicladores computadorizados realizam as variações térmicas cíclicas de maneira rápida e precisa. Para ocorrer a amplificação de um trecho de DNA específico é necessário um oligonucleotídeo complementar a este trecho, para garantir o pareamento e iniciar a replicação. Esse oligonucleotídeo é denominado de *Primer* [43].

Só haverá amplificação do DNA se houver hibridização do *primer*, servindo como um modelo do trecho do DNA a ser procurado. Devido a isso, é importante que seja projetado um bom *primer* a ser utilizado, para que aquela sequência somente seja encontrada no microrganismo alvo. Essa é a técnica que utiliza uma bactéria resistente a altas temperaturas e de onde é extraído uma DNA polimerase Taq polimerase [18].

Os fatores que interferem na eficácia da PCR são a qualidade do DNA, que deve estar livre de contaminantes, a escolha correta dos *primers* para a amplificação do material, assim como o emprego de métodos de extração adequados, especialmente para amostras paucibacilares [42] [40].

Apesar da PCR ser um método de diagnóstico promissor, especialmente nos locais com índices de tuberculose elevados, ela ainda apresenta limitações não só relativa à padronização, chance de contaminação da amostra e à extração do DNA, como também ao custo, complexidade, qualidade e armazenamento dos reagentes [44].

Uma variação da PCR tradicional é o método *Nested PCR*[45]. Inicialmente o produto genômico é amplificado com uma sequência maior do que a desejada, e depois, a partir do produto desse processo, é extraída a sequência alvo para passar pelo mesmo processo. Essa técnica é útil quando as impurezas que inibem a amplificação do DNA estão presentes nas amostras ou quando a quantidade de DNA do

agente que se deseja detectar é muito pequena. A vantagem que pode ser vista é o aumento da Especificidade e Sensibilidade, porém o tempo de realização e o custo são maiores [46].

### **2.2.6 Anergia**

O diagnóstico de um animal com suspeita de contaminação por tuberculose bovina é feito com a combinação de um ou mais exames mencionados anteriormente. Dependendo do estágio da doença no animal e da quantidade de possíveis animais infectados no rebanho, um diagnóstico será suficiente ou deverão ser feitos outros testes dentro das etapas explicitadas [47] [48].

Existem porém animais infectados que podem não reagir aos exames de testes cervicais (TCS e TCC), fazendo com que o diagnóstico seja falho. Esses animais são chamado de anérgicos.

A anergia se caracteriza pela ausência de resposta imunológica e pode ser congênita ou adquirida. Tal estado também é definido, especificamente ao problema da Tuberculose, como a falha de um animal com sinais visíveis de Tuberculose Bovina em exibir resposta hipersensitiva ao teste da tuberculina [49]. Isso em geral acontece no estágio avançado da doença, que pode causar a falta de resposta por parte do animal devido à alta carga bacteriana, pois os animais acabam tendo um diminuição na sensibilidade alérgica [47] [48].

Para esses casos, os exames sorológicos como o ELISA se mostram vantajosos, pois possibilitam chegar a um diagnóstico mais conclusivo. A ocorrência de animais anérgicos coloca em risco os programas nacionais de erradicação da tuberculose bovina [47]. As ocorrências de animais anérgicos podem variar de acordo com a região e rebanho, como em [50] que acabou apresentando aproximadamente 1% de animais anérgicos de um total de 560 animais, e também em [51] onde foram encontrados 2 animais anérgicos em um total de 360 bovinos.

## **2.3 Impacto Econômico**

A tuberculose bovina está presente em todo o território nacional e é responsável por perdas econômicas importantes na pecuária brasileira [21]. A doença abrange diversos estados e regiões do Brasil, como por exemplo os estados de Mato Grosso do Sul (MS) [52], Rio de Janeiro (RJ) [53], Paraíba (PB) [54], como também São Paulo (SP), Minas Gerais (MG) e Goiás (GO) [55]. Órgãos oficiais estimam que a prevalência da tuberculose bovina esteja próximo de 1% do total de bovinos no Brasil [26].

Os prejuízos causados são principalmente a diminuição da produção de leite,

eliminação de animais, condenação de carcaças no abate e na comercialização dos produtos derivados [20]. Um animal tuberculoso pode apresentar de 10% a 25% de queda na capacidade produtiva [3] e tal queda é vista tanto na produção de leite como também na produção de carne.

No Brasil não existem dados sobre o impacto da doença na produção de leite, porém um estudo, realizado na Argentina em 1988, mostra que as perdas na produção do leite por animais tuberculosos atingiram 18% [56] [36].

Os bovinos tuberculosos são uma fonte de infecção para outros animais e também para o homem, gerando assim um custo relacionado ao tratamento da doença nos humanos como na prevenção da tuberculose em outras espécies.

Um exemplo desses impactos no Brasil é a grande concorrência que o país vem enfrentando frente ao mercado internacional, provocada pelas exigências sanitárias cada vez mais rigorosas impostas pelos principais países importadores de derivados de origem animal. O país acaba comercializando com mercados com menos restrições sanitárias e que comprem os produtos por um preço menor [57].

## 2.4 Programa de Controle e Erradicação

A doença está amplamente distribuída em todo o mundo, sendo alvo de programas de erradicação em diversos países. No Brasil, a importância do controle da doença levou o Ministério da Agricultura, Pecuária e Abastecimento, a instituir, em 2001, o *Programa Nacional de Controle e Erradicação da Brucelose e da Tuberculose* (PNCEBT), que é um programa de adesão voluntária que regulamenta as estratégias de controle e visa reduzir a prevalência da doença no país [21] [26].

O programa prioriza o teste de tuberculinização como diagnóstico, e se baseia na periodicidade desses testes para o controle da doença, e conseqüente sacrifício dos animais infectados. Esses testes devem ser feitos anualmente em áreas de produção de leite, e com um intervalo um pouco maior em áreas de gado de corte, já que existem exames que serão realizados no abate e possibilitam a observação de lesões [22] [27] [3].

Outros métodos que auxiliam na identificação da doença nos rebanhos são as inspeções sanitárias nos produtos de origem animal e os exames de necropsia feitos por veterinários nos animais suspeitos de tuberculose. Outra medida importante é a inspeção feita por parte dos frigoríficos responsáveis pelo abate do animal tuberculoso, onde uma análise mais detalhada e cuidadosa pode mostrar a presença da tuberculose no bovino e conseqüentemente na fazenda de origem [21].

Outra medida adotada no PNCEBT é a certificação, válida por 12 meses, de estabelecimentos livre de tuberculose emitidos pela Delegacia Federal da Agricultura. Para obter o certificado é necessário que sejam realizados três testes tuberculínicos

consecutivos retornando resultados negativos de todo o rebanho [3]. Esses testes devem ser realizados em um intervalo de 90 a 120 dias entre o primeiro e o segundo teste, e de 180 a 240 dias entre o segundo e o terceiro [3].

O objetivo do PNCEBT é estimular os criadores a buscarem a certificação com a justificativa que tal certificado agregue valor aos seus produtos. Porém, os principais impedimentos para a adesão ao certificado são as próprias medidas sanitárias propostas e o fato do custo dos testes ser arcado pelos próprios criadores. Além disso, em resultado positivo do teste, o prejuízo da perda do animal, redução de produção, infertilidade, entre outros, seria arcado pelos criadores e produtores [3] [36].

# Capítulo 3

## Técnica de Aprendizado de Máquina

O campo de Aprendizado de Máquina (*Machine Learning*) é uma ramificação do campo de Inteligência Artificial e da computação. A principal característica dessa área de estudo é a capacidade de aprendizado por um sistema apenas a partir de dados, que representam o problema a ser aprendido, sem a necessidade de uma programação específica. As aplicações para este campo de estudo estão por exemplo, em problemas de regressão ou aproximação de funções, classificação, reconhecimento de padrões, otimização e análise temporal [58].

Existem diferentes abordagens ou técnicas de Aprendizado de Máquina que são utilizadas para diferentes problemas e aplicações, como por exemplo: SVM (*Support Vector Machines*) [6], Redes Neurais Artificiais [59], Algoritmo Genético [58], Clusterização e Árvores de decisão (*Decision Trees*) [60].

A técnica utilizada neste estudo foi a Rede Neural Supervisionada e por isso o foco deste capítulo está sobre esta técnica de aprendizado de máquina. A Seção 3.1 apresenta os conceitos acerca das Redes Neurais Artificiais, além dos conceitos referentes ao paradigma supervisionado de treinamento. A Seção 3.2 aborda métodos estatísticos que foram considerados neste estudo.

### 3.1 Redes Neurais Artificiais

Redes Neurais Artificiais (RNA) são sistemas computacionais e uma de diversas técnicas de Aprendizado de máquina, sendo seu conceito inspirado em estruturas biológicas, mais precisamente no cérebro humano, e em seu modo de aprendizado [7]. HAYKIN *et al.* [6] define uma Rede Neural Artificial como “um processador massivamente paralelo e distribuído, feito de unidades simples de processamento e com a propensão natural em armazenar e utilizar o conhecimento”.

Suas aplicações podem ser feitas em problemas de classificação, análise temporal, estimação de valores e problemas de regressão. Exemplos práticos já apresentaram bons resultados e performance com seu uso, como por exemplo em problemas de detecção de partículas [13], reconhecimento de faces humanas [12] e ao diagnóstico de Tuberculose Pulmonar em humanos [61].

Nesta seção são apresentados os principais conceitos sobre Redes Neurais Artificiais, com ênfase em seu caso de aprendizado Supervisionado.

### 3.1.1 Tipos de Aprendizado

Uma das características mais interessantes de uma RNA é a sua capacidade de aprendizado e este é o primeiro conceito considerado para elas. O conhecimento ou o aprendizado de uma RNA é obtido a partir dos dados que lhe são apresentados e que representam o conhecimento a ser adquirido. O processo pelo qual a rede neural aprende a partir de dados disponíveis é chamado de “Treinamento” e este processo pode ser feito seguindo diferentes paradigmas de aprendizado. Os tipos de aprendizado considerados são: Supervisionado, Não Supervisionado e por Reforço [58] [62].

O primeiro paradigma é o chamado “Aprendizado Supervisionado” e este consiste na representação de um conhecimento já estabelecido e conhecido nos dados. No caso, os dados usados no aprendizado já possuem um valor alvo ou uma resposta definida. Neste tipo de treinamento, a capacidade da Rede Neural está em conseguir aproximar a resposta que ela retorna com o valor alvo que antes lhe foi apresentado. Os dados são formados por pares, onde cada par apresenta valores de entrada que representam a sua saída, associada com seu valor alvo.

O segundo caso considerado é o paradigma de aprendizado “Não Supervisionado”, que consiste basicamente do caso em que os dados apresentados à rede neural não possuem um valor alvo definido e portanto, a rede não possui um parâmetro para aproximar os seus valores de saída. A abordagem que este paradigma então define é de se buscar similaridades entre os dados apresentados para que estes sejam então agrupados ou categorizados.

O terceiro caso, chamado paradigma de aprendizado por “Reforço” e ele pode ser considerado como um caso intermediário entre o caso Supervisionado e o Não Supervisionado [58]. Isso ocorre porque neste paradigma de aprendizado os dados utilizados pela Rede Neural não possuem um valor alvo definido, mas possuem a informação de se a saída obtida pela rede está correta ou não, porém sem indicar o valor a ser aproximado [58]. Uma outra abordagem para o aprendizado por reforço é atribuir a cada par de entrada e saída, um valor associando a qualidade desta saída e medindo o quão boa é o valor gerado [62].

### 3.1.2 Neurônio Biológico e Neurônio Artificial

Nos animais, o aprendizado ocorre através do cérebro e com o maior entendimento de sua estrutura e funcionamento, é possível copiar características suas para utilizar na aprendizagem de máquina. O cérebro é um sistema poderoso e complexo, capaz lidar com ruídos, com dados inconsistentes, e produzir respostas que normalmente são corretas para dados tanto de altas quanto de baixas dimensões muito rapidamente [58].

O cérebro é composto por aproximadamente 85 bilhões de células nervosas chamadas “Neurônios”, que podem ser consideradas as unidades básicas do mesmo, como também de todo o sistema nervoso. Em analogia entre a computação e a área biológica, tais células podem ser como consideradas como os processadores do cérebro [58]. A estrutura do neurônio é relacionada na Figura 3.1, seguida da identificação e descrição de alguns de seus componentes [1].

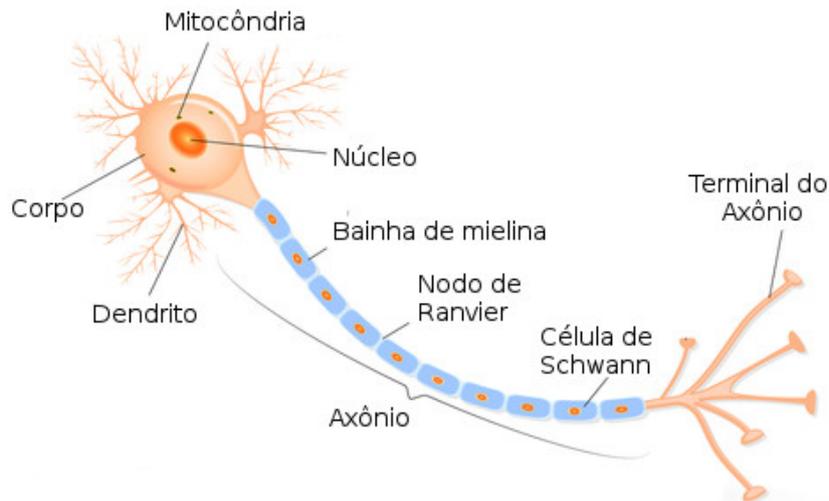


Figura 3.1: Estrutura do Neurônio Biológico [1]

- **Corpo celular:** Também chamado de “Soma” ou “Pericário”, é a principal parte do Neurônio e onde estão todos os componentes necessários de uma célula. É o componente onde ocorrem os processos metabólicos e de saem as estruturas dos Dendritos e do Axônio.
- **Axônio:** É semelhante a um cabo e transporta os impulsos recebidos à outras células, como as musculares, glandulares ou mesmo outros neurônios. Ele pode ser coberto por uma fina camada de gordura chamada Mielina, que funciona como um isolamento elétrico e ajuda a acelerar a transmissão do impulso. É o responsável por transmitir adiante a informação gerada pelo Núcleo.
- **Dendritos:** São os numerosos prolongamentos e saem do Corpo Celular. São

os responsáveis por receber os estímulos do ambiente, células epiteliais sensoriais e outros neurônios.

- **Sinapse:** É a conexão entre o Dendrito e o Axônio. É a ligação entre dois Neurônios distintos e a responsável pela efetiva passagem dos impulsos elétricos.

A informação no cérebro é transmitida pelos neurônios no formato de impulsos elétricos, que são recebidos pelos dendritos do Corpo Celular do Neurônio. Essa recepção dos impulsos ocorre somente quando os dendritos do corpo celular estão conectados com o terminal do axônio de outro neurônio, configurando assim uma ligação sináptica. A informação é transmitida pela sinapse graças a uma variedade de substâncias químicas chamadas “Neurotransmissores”. Com os impulsos recebidos, caso estes ultrapassem um limiar interno do próprio neurônio, este também transmite um novo impulso elétrico. Esse impulso é transportado através de seu axônio, e caso haja uma ligação sináptica em seus terminais, o mesmo é transmitido adiante até outras células cerebrais [1].

Cada neurônio pode ser visto como um processador separado e simplificado, com a única tarefa de processar a decisão de enviar ou não o impulso. Isso faz do cérebro um computador massivamente paralelo composto de  $85^9$  processadores. O cérebro tem a capacidade de se auto organizar fazendo com que seus neurônios consigam processar reconhecimento de padrões, percepção de sentidos, controle motor e aprendizagem [58].

Baseado nesses conceitos e no funcionamento da estrutura biológica, McCulloch e Pitts (1943) propuseram um modelo matemático inspirado nos componentes do neurônio biológico, sendo este chamado de Neurônio Artificial. Eles modelaram o neurônio com as características a seguir [7]:

- **Entrada:** O neurônio artificial recebe várias entradas, simulando os impulsos elétricos captados pelos dendritos. As entradas são os dados que serão processados pelo Neurônio artificial para a geração da saída.
- **Pesos:** Cada entrada do neurônio possui um peso associado, para simular o funcionamento das sinapses. Cada peso simula a importância que um sinal tem para o neurônio e sua força
- **Função de Ativação:** É a função do limiar onde é decidido se o neurônio transmite o impulso ou não a partir da entrada corrente.
- **Saída:** Assim como o axônio, o neurônio transmite uma saída relacionada ao seu conjunto de entradas. O valor da saída é definido de acordo com a função de ativação usada.

Relacionando as estruturas do neurônio biológico com o neurônio artificial é possível identificar os conceitos biológicos considerados por McCulloch e Pitts na elaboração de seu modelo. As entradas que o neurônio artificial recebe é uma abstração para os impulsos elétricos recebidos pelos dendritos do neurônio biológico e são com essas entradas que o neurônio artificial irá gerar sua saída. O conceito de “Pesos” está relacionado ao conceito de sinapses nervosas, e ele atribui para cada entrada uma importância ou força que é considerado pelo neurônio artificial. No neurônio biológico esse fato também é visto, uma vez que a quantidade de neurotransmissores em sinapse é quem define a relevância do impulso elétrico que é propagado ao próximo neurônio. A chamada Função de Ativação é análoga ao limiar do neurônio biológico, e é o responsável por transformar as entradas no valor de saída do neurônio artificial. Por fim, a saída do neurônio artificial equivale ao impulso elétrico gerado pelo neurônio biológico e que é transmitido adiante aos demais a ele ligados.

A Figura 3.2 mostra a representação desse modelo de Neurônio Artificial. Nele, uma série de entradas  $x_1, x_2, \dots, x_m$  são apresentadas ao neurônio. Cada sinal de entrada é multiplicado por um peso associado  $w_1, w_2, \dots, w_m$ , onde cada peso corresponde à força de um impulso elétrico. Cada entrada é então multiplicada por seu peso correspondente e todos esses valores são somados. A função de ativação se limita a considerar se esse resultado é maior ou igual ao seu limiar, e então a função é ativada [58] [7]. As estruturas do neurônio artificial responsáveis pelo somatório dos valores da combinação das entradas e dos pesos, tal como a estrutura que aplica a função de ativação, podem ser comparadas ao Corpo Celular do neurônio biológico.

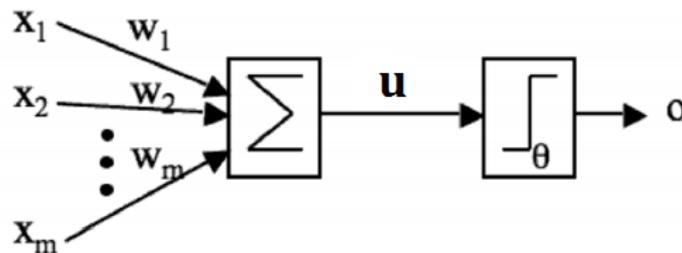


Figura 3.2: Modelo de Neurônio Artificial de McCulloch e Pitts

O modelo de McCulloch e Pitts é simples, e com o tempo foi aperfeiçoado. A primeira mudança foi o acréscimo do componente *BIAS* que é somado junto aos valores das entradas multiplicadas pelos pesos e que também possui um peso associado. O valor atribuído ao *BIAS* é igual à 1 e com isso, o sinal vindo deste elemento depende apenas do valor de seu peso sináptico. Outra mudança foi na função de ativação, onde foram introduzidos outros tipos de funções. No primeiro modelo de neurônio artificial, a função de ativação utilizada era uma função binária, onde apenas era

enviado o valor 0 ou 1. Com as novas funções, os valores podem variar de acordo com a função escolhida para o neurônio. São usualmente utilizadas as funções a seguir [6]:

1. **Função Limiar:**

$$f(u) = \begin{cases} 1, & u \geq 1 \\ 0, & u < 1 \end{cases}$$

2. **Função Linear:**

$$f(u) = k * u$$

3. **Função Sigmoide:** Também chamada de Função Logística, é uma função contínua que permite a transição gradual entre os dois estados (0 e 1).

$$f(u) = \frac{1}{1 + e^{-u}}$$

4. **Função tangente hiperbólica:** É uma variante da Função o Sigmoide. Seus valores variam entre (-1,1).

$$f(u) = \frac{1 - e^{-u}}{1 + e^{-u}}$$

A Figura 3.3 relaciona o formato das funções de ativação expostas.

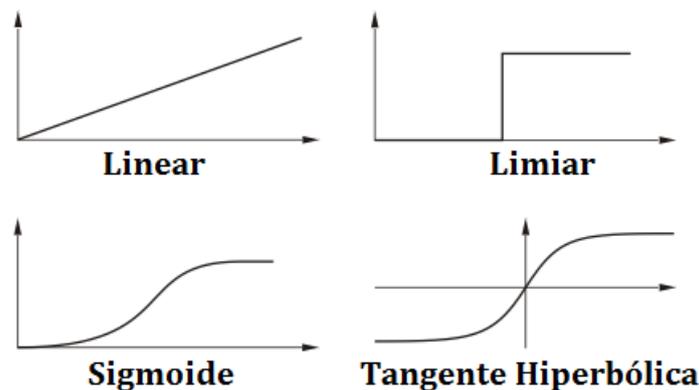


Figura 3.3: Formato das Funções de Ativação consideradas

Com as alterações propostas, o neurônio artificial apresenta o formato relacionado na Figura 3.4. Nela vê-se a adição do componente *BIAS* e da definição de uma função de ativação genérica, podendo ser qualquer função especificada para o neurônio.

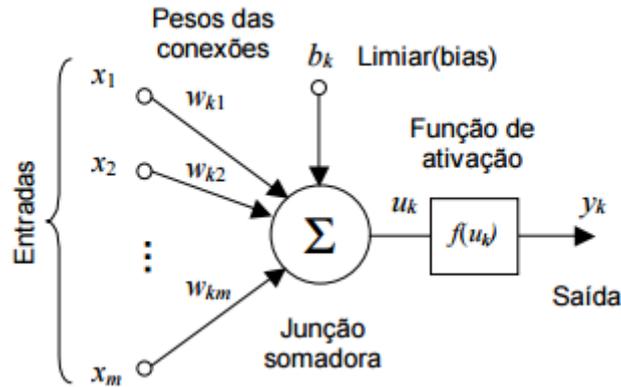


Figura 3.4: Modelo de Neurônio Artificial

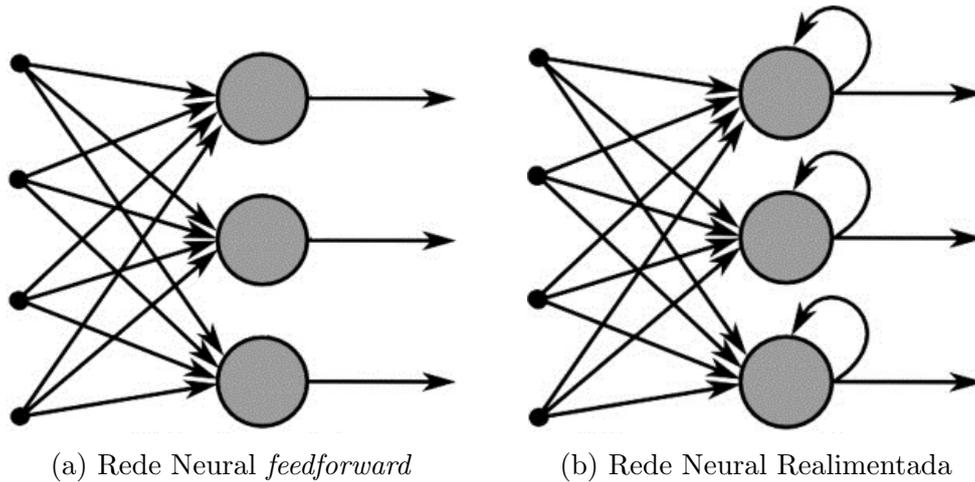
As funções Sigmoide e Tangente Hiperbólica são importantes funções de ativação por introduzirem um fator de não linearidade às redes neurais. As redes neurais possuem a capacidade de aproximar funções matemáticas, porém, sem o uso de funções não lineares elas são apenas capazes de aproximar combinações lineares de suas entradas. Esta capacidade de aproximação de qualquer função linear ou não linear é dada pelo Teorema da Aproximação Universal [63]. Outro ganho com o uso destas funções é o computacional. Existem diferentes métodos para o aprendizado das redes e métodos que usam os conceitos de derivadas apresentam ganhos quanto ao tempo de processamento, uma vez que as funções sigmoide e tangente hiperbólica tem suas derivadas computacionalmente calculadas mais facilmente.

### 3.1.3 Arquitetura das Redes Neurais Artificiais

Com a definição do neurônio artificial, o próximo fator abordado para as redes neurais são suas topologias e arquiteturas. A Rede Neural é basicamente composta por neurônios artificiais interconectados por sinapses, havendo ainda as sinapses que conectam as entradas das redes com seus neurônios.

A arquitetura e a topologia de uma RNA pode ser definida por fatores como, os tipos de conexão que ela possui, o número de neurônios utilizados para treinamento, número de camadas existentes e a função de ativação. Essa definição é importante para que cada arquitetura possa resolver adequadamente um dado problema.

Uma organização comum para os neurônios é a sua separação por camadas, definidas pelo projetista da Rede Neural e influenciado pelo problema aplicado. Cada camada é composta por neurônios artificiais. Para essa organização em camadas, existem dois tipos de RNAs, onde o que as difere é o tipo de conexão das sinapses que é utilizada. Esses dois tipos de rede são as chamadas “*feedforward*” e as chamadas “Realimentadas”. A Figura 3.5 relaciona os dois tipos de RNA citados.



(a) Rede Neural *feedforward*

(b) Rede Neural Realimentada

Figura 3.5: Redes Neurais e seus tipos de conexão

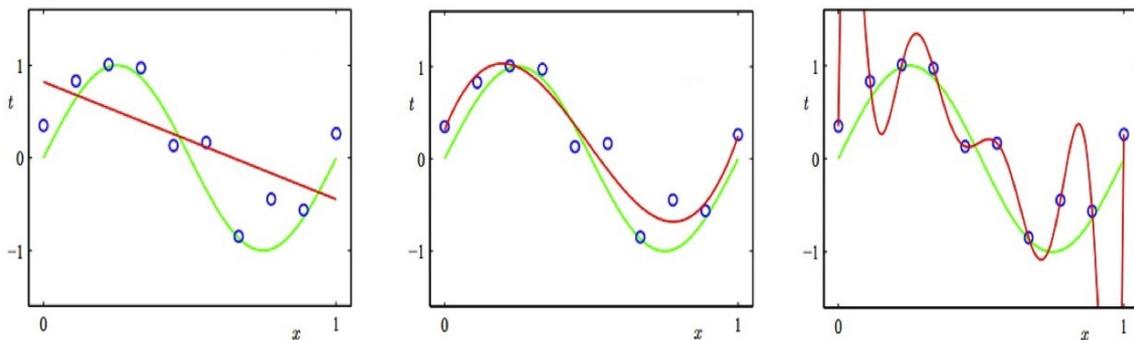
As redes neurais do tipo *feedforward* são redes em que as saídas do neurônio de uma camada não possuem nenhuma conexão com neurônios de camadas anteriores ou neurônios da mesma camada (com ele próprio incluído), elas se conectam apenas com a seguinte. Todas as sinapses portanto, têm a mesma direção, partindo da camada de entrada para a de saída. A Figura 3.5(a) apresenta um exemplo de RNA *feedforward*, onde se vê a partir da esquerda, que as entradas são conectadas aos neurônios da RNA e que apenas “para frente” (*forward*) que a saída do neurônio é enviado, no caso do exemplo, como saída da rede.

Já nas redes com conexões realimentadas, os neurônios podem utilizar suas saídas para realimentar as entradas, outros neurônios da mesma camada ou anteriores, como também a si próprios. Elas ainda podem admitir memória a curto prazo, aproximando-as mais das características dos cérebros. São mais utilizadas para resolver problemas que envolvem processamento temporal [64]. A Figura 3.5(b) apresenta um exemplo de RNA realimentada e nela é visto que os neurônios da rede apresentam a si mesmos as saídas geradas anteriormente como uma nova entrada aos neurônios.

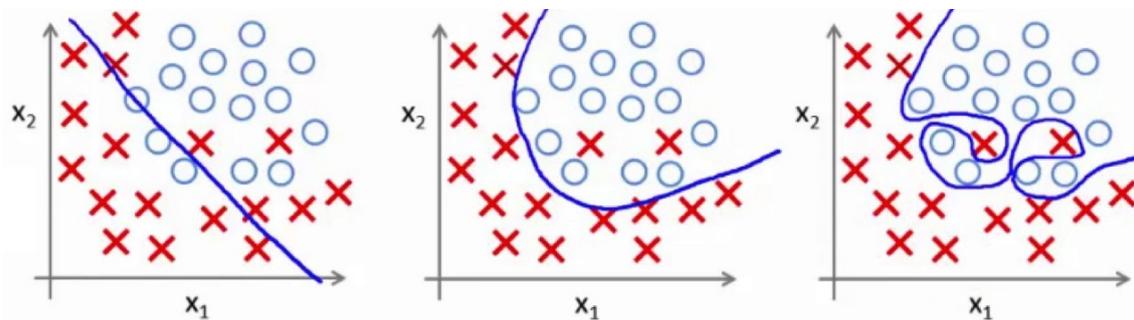
Outro fator importante na topologia de uma rede neural é o número de neurônios utilizados, sendo esta quantidade definida de acordo com o problema e o domínio estudado. É um fator que deve ser escolhido com cuidado, uma vez que influencia diretamente na capacidade de aprendizagem e generalização da rede. Duas situações indesejadas podem ser geradas pela escolha de uma quantidade escassa ou excessiva de neurônios, e elas são respectivamente, os problemas de *Underfitting* e *Overfitting*, diretamente ligados à capacidade de generalização do sistema.

A habilidade de generalizar refere-se à capacidade de um sistema de Aprendizado de Máquina em, a partir de sua experiência ou do aprendizado obtido a partir dos dados apresentados, conseguir apresentar as saídas corretamente para entradas novas e ainda não vistas por ele (“novas experiências”) [59]. Os dados usados para o

aprendizado da rede ou de qualquer sistema inteligente, devem ser idealmente compostos por amostras representativas para todo o espaço de ocorrências do problema, para que assim, o modelo possa aprender sobre todos os casos presentes no problema e melhor generalizá-los. A Figura 3.6 apresenta exemplos de *Underfitting*, uma boa generalização e *Overfitting* para problemas hipotéticos de regressão (Figura 3.6(a)) e classificação (Figura 3.6(b)). Para o exemplo de regressão, a linha destacada na cor “verde” representa a função alvo que o sistema de aprendizado deseja buscar, enquanto a linha destacada em “vermelho” trás a função que foi gerada na saída do sistema. No caso de classificação, deseja-se apenas separar as amostras destacadas de maneira eficiente.



(a) Casos de Regressão



(b) Casos de Classificação

Figura 3.6: Exemplos de *Underfitting*, de boa generalização e *Overfitting*

O problema de *Underfitting* é dado pela incapacidade do modelo de aprendizado em generalizar o problema lhe foi apresentado, tanto para os dados que lhe são passados para seu treinamento como para dados nunca vistos pelo modelo e que não fazem parte deste processo [65] [66]. A pouca generalização por parte do modelo significa que ele não consegue encontrar a relação entre os valores de entradas e os valores alvos (saídas desejadas). Uma situação de *Underfitting* é usualmente observada apenas a partir dos dados que a rede neural dispõe para o seu aprendizado, onde no caso ela apresenta grandes erros comparados aos seus valores alvo. Para as RNAs, uma pequena quantidade de neurônios pode causar esse tipo de problema.

O gráfico mais à esquerda da Figura 3.6(a) mostra um exemplo de *Underfitting* em um caso de regressão e nele pode-se ver que o comportamento desejado é o de uma função não linear, porém o modelo apenas consegue aproximar uma função linear, apresentando um resultado longe do desejado. Para o exemplo na classificação, na Figura 3.6(b) mais à esquerda, é visto que o modelo gera um separador entre as classes com baixa eficiência, classificando erradamente amostras de ambas as classes.

Já a situação de *Overfitting* é dada pela perda de generalização do problema porém, pela razão inversa ao *Underfitting*, pois consegue classificar os dados que lhe são dados para o aprendizado muito bem, mas de modo que o modelo se torne muito específico para estes dados [65]. O *Overfitting* é visto ainda quando o fenômeno de acertar os dados de treinamento não indica mais que o modelo seja capaz de generalizar amostras fora deste grupo [62]. Em uma RNA, o que se vê nesse caso é que ela irá aprender baseado nos detalhes dos dados usados, inclusive nos erros e ruídos que possuam, e não apenas nas características mais gerais e comuns à maioria. O uso de uma grande quantidade de neurônios pode levar à este caso. Os gráficos mais à direita nas Figuras 3.6(a) e 3.6(b) relacionam os casos de *Overfitting* para os casos de regressão e classificação, respectivamente. Para a regressão, o que se vê é que o modelo retorna uma função que consegue passar exatamente por todos os pontos dos dados que lhe foram passados, porém esta função possui um grau de complexidade maior do que a função alvo. Na classificação, observa-se que o separador separa as amostras perfeitamente, porém sem considerar que duas destas amostras eram provavelmente dados com erros e nesse caso, o separador levou esse erro em consideração para gerar sua saída.

Por fim, uma situação de equilíbrio é alcançada quando não é usada nem uma pequena e nem uma grande quantidade de neurônios nas RNAs. Os gráficos do meio nas Figuras 3.6(a) e 3.6(b) mostram exemplos de saídas com boa generalização para os exemplos considerados.

Quanto às camadas de neurônios das RNAs, estas podem possuir uma ou várias camadas. Redes com uma camada possuem apenas a camada as entradas e a camada de processamento, que já é a camada de saída, enquanto as redes Multicamadas possuem também as camadas escondidas, entre a entrada e a de saída. A Figura 3.7 relaciona um exemplos de RNA com uma e duas camadas de neurônios.

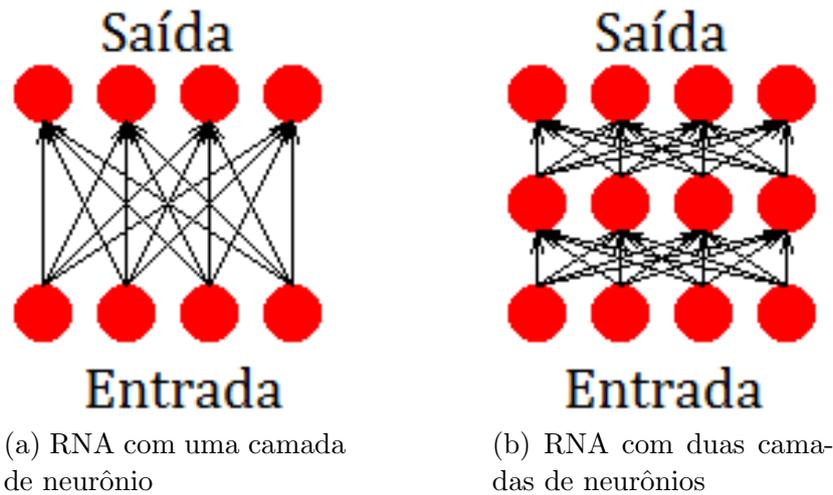


Figura 3.7: Exemplos de Redes Neurais com uma e duas camadas de neurônios

A seguir, é abordado especificamente as Redes Multicamadas Perceptron e suas características.

### 3.1.4 Redes Multicamadas Perceptron

Dentre os principais tipo de rede Feedforward, destaca-se as Redes Multicamadas Perceptron ou MLPs (*Multi Layer Perceptrons*). Sua arquitetura é definida pela utilização de duas ou mais camadas de neurônios e conexões do tipo *feedforward*. Ela é composta por uma entrada, onde são imputados os dados, uma ou mais camadas intermediárias de neurônios, também chamadas de camadas “internas” ou “escondidas”, e uma camada de saída. Em todas as camadas existe o componente “bias”, com seu valor sendo constante e igual a 1. As MLPs possuem conexões do tipo *feedforward*, logo o sinal é direcionado da entrada à camada de saída, sendo propagado através de cada uma em sentido progressivo [6] [7]. A Figura 3.8 apresenta o formato geral que uma MLP possui.

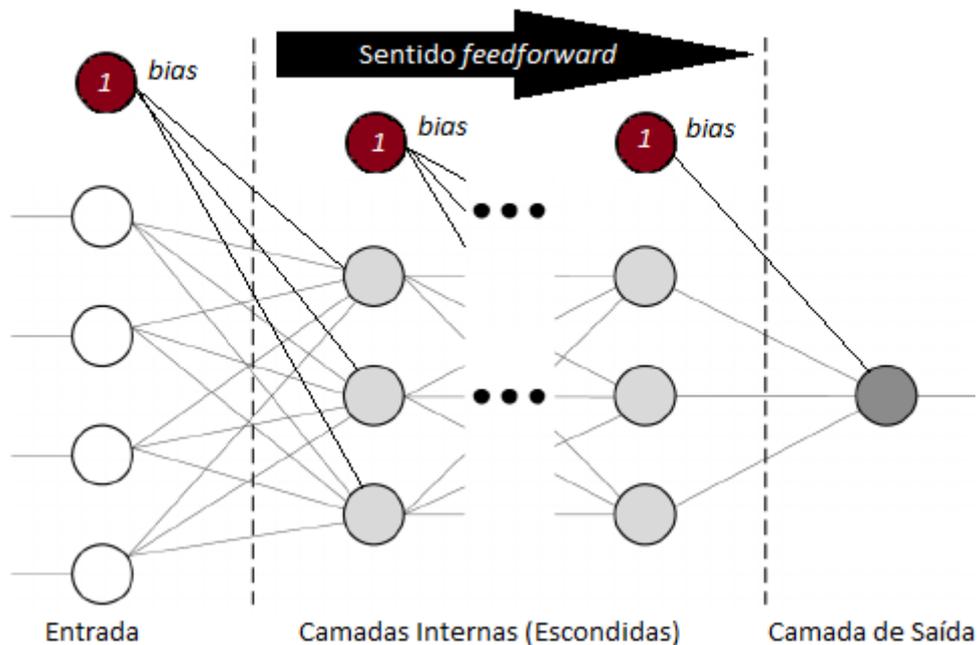


Figura 3.8: Formato de uma rede MLP

O primeiro fator a ser considerado é o porque do uso de das redes MLPs em detrimento das RNAs com apenas uma camada de neurônios, e este se dá pela capacidade da MLP em resolver problemas mais complexos que não são aproximados pelas rede mais simples. Uma rede com apenas uma camada possui a limitação de resolver apenas problemas lineares e não conseguir usar informações de ordem superior ou não linear [62]. Já as redes MLPs, por possuírem mais camadas e aliadas a isso também poderem usar funções de ativação não lineares, conseguem abordar os problemas de ordem superior tal como os lineares [67]. Um exemplo foi visto na Figura 3.6, onde os casos de *Underfitting* nela destacados, mostram também o comportamento de RNAs com capacidade apenas de resolução de problemas de baixa ordem, enquanto os resultados ideais (de “*fitting*”) mostram exemplos da capacidade de uma rede para problemas de ordem superior.

Nas camadas intermediárias ocorre a extração de características, e em geral são utilizadas como funções de ativação as funções Logarítmicas ou Sigmoides e suas variantes (Ex: Tangente Hiperbólica) [68]. O uso destas funções de ativação se mostram importantes pela introdução não da linearidade às redes MLP e a capacidade de resolver problemas de ordem superior, além dos de baixa ordem.

Nas Redes Neurais, seu conhecimento está armazenado em seus pesos sinápticos e sua fase de aprendizado é definida pelo ajuste destes, além de poderem ser gerados inicialmente de forma aleatória e serem ajustados ao longo do treinamento da rede, a partir de uma taxa de aprendizagem ou taxa de momento que influenciará no tempo de processamento [6] [68]. O tipo de aprendizado utilizado em uma MLP é o aprendizado supervisionado e por isso, em seu treinamento busca-se obter o

valor alvo definido para cada entrada que a rede recebe. No entanto, esses dados de entrada podem possuir ruídos, erros estatísticos, ou mesmo o problema apresentar uma complexidade maior do que capacidade de resolução da rede, fazendo com que ela então considere um erro sobre sua saída, em relação ao valor objetivo. Assim, o treinamento das MLPs busca em geral minimizar o erro associado às suas saídas e esta medida de erro pode ser usada como critério de avaliação do desempenho de uma rede.

Outra característica importante de uma MLP é o critério utilizado para saber em que momento interromper o treinamento, uma vez que nesta etapa podem ser encontradas as situações de *Underfitting* e *Overfitting*. Critérios de parada podem ser definidos simplesmente como um número estimado de épocas necessárias para se alcançar o resultado, ou serem determinados pelos erros de saída das redes, sendo esta segunda abordagem muito utilizada, além de estar ligada ao modo de avaliação de uma rede neural. Para se avaliar o desempenho de uma MLP e também determinar o seu momento de parada são definidos três grupos e suas funções sobre os dados utilizados, chamados de Grupos de Treinamento, Validação e Teste [6].

O conjunto de treinamento é a parcela dos dados que é atribuída apenas para o aprendizado do sistema inteligente, sendo a partir apenas desses dados que o modelo irá ajustar os valores de seus pesos. Por ser um conjunto importante e que irá definir o funcionamento da MLP, idealmente este conjunto deve ser o mais representativo possível em relação ao problema, para que o modelo ensinado possa cobrir a maior variedade possível de ocorrências do mesmo.

Já o conjunto de Validação não é usado no aprendizado da MLP, mas geralmente para a avaliação do treinamento e determinação de parada do mesmo. Como visto, o *Overfitting* pode ser dado por uma boa performance do modelo para seu grupo de treinamento mas sem nenhuma capacidade de generalização ao encontrar amostras fora destes dados de treinamento e sendo assim, o grupo de validação é usado como esses dados fora da amostra de treinamento. O conjunto de Teste é similar ao grupo de Validação, porém este é usado apenas após toda etapa de treinamento ter sido feita e tal grupo tem como objetivo avaliar a performance, segundo a medida de erro escolhida, do sistema já terminado.

A Figura 3.9 relaciona um exemplo para a determinação de parada de treinamento de uma MLP usando os erros do grupo de treinamento e validação.

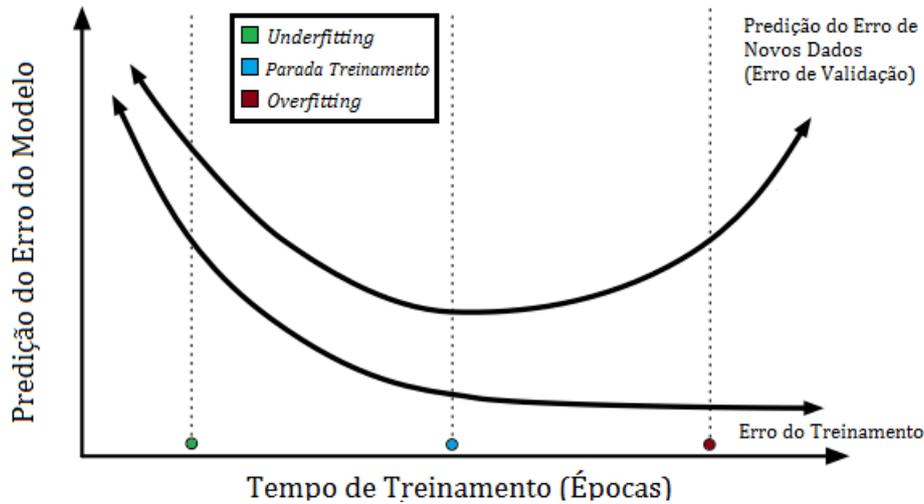


Figura 3.9: Exemplo de Erros em treinamento de rede MLP

O exemplo usa uma métrica genérica para o erro e o que se vê é que na medida que as épocas progridem, o erro do conjunto de treinamento cai, enquanto que para o erro do grupo de validação inicia apresentando uma queda e a partir de certo ponto aumenta. O ponto indicado na cor vermelha indica uma situação de *Overfitting* no treinamento, onde o erro para o conjunto de treinamento é baixo e o erro do grupo de validação é alto, tendo este crescido depois de certo ponto. Já o ponto indicado na cor verde indica uma situação de *Underfitting*, com ambos os erros altos e ainda decaindo. O ponto destacado na cor azul é o ponto que indica que o treinamento deve ser encerrado, pois a partir deste momento o erro de validação começa a aumentar, indicando que o modelo está perdendo capacidade de generalização do problema.

É ainda importante destacar que a forma como é feita a divisão dos dados nesses três grupos tem relação direta com os erros medidos em cada, uma vez que amostras mais ou menos representativas podem estar ou não neles. O problema analisado e a quantidade de dados disponíveis são fatores que podem influenciar nessa divisão dos dados [58].

O modo como é feito o ajuste dos pesos é também um fator muito importante para definir a capacidade de generalização de uma MLP e para esta tarefa existem dois algoritmos que serão considerados a seguir, os algoritmos *Backpropagation* e *Resilient Backpropagation* [68].

### 3.1.5 Algoritmo *Backpropagation*

Após esboços feitos em anos anteriores por outros pesquisadores, Rumelhart, Hinton e Williams apresentaram em 1986 uma descrição clara e concisa sobre o algoritmo de *Backpropagation* [7]. Durante muito tempo não havia um algoritmo para treinar redes multicamadas e o desenvolvimento do algoritmo se deu devido ao retorno do interesse em redes neurais artificiais, e expandiu dramaticamente a gama

de problemas que podem ser resolvidos pelas RNAs. Ele é um algoritmo utilizado de treinamento com base matemática [7] e também um algoritmo de aprendizagem supervisionada específico para redes *feedforward* e MLPs, sendo amplamente usado e muito popular [69].

O diferencial do *Backpropagation* ocorre em dois passos, onde no primeiro, a rede é iniciada com pesos aleatórios dentro de um intervalo estabelecido, e são inseridos os dados na camada de entrada. Essa informação é propagada através da rede, por cada camada, até gerar os resultados na camada de saída. No segundo passo, a saída retornada pela rede é comparada com o valor objetivo e então calcula-se o erro gerado pela amostra. Tal erro é então retropropagado a partir da camada de saída para as camadas anteriores, realizando o efeito de *Backpropagation*, ou seja, a propagação de trás para frente, e assim os pesos são recalculados e ajustados baseados nesse erro. Esse processo se repete até que certos critérios sejam atendidos, como por exemplo certo número de iterações transcorridas ou um valor aceitável de erro [68].

Após o término do treinamento, a rede pode ser utilizada para propagação e teste de outros dados, sendo usada apenas no sentido tradicional (*feedforward*). Os dados são inseridos pela camada de entrada, processados pela rede e o resultado é gerado em sua camada de saída. Nesta fase porém, o erro calculado não é retropropagado na rede e não há reajuste dos pesos.

O *Backpropagation* é baseado no gradiente decrescente, o que faz com que ele não necessariamente encontre um mínimo global e possa ficar parado em mínimos locais [68]. O erro em um problema de *Machine Learning* pode ser representado por uma malha e os pontos mais baixos, como um vale ou depressão, são os valores de mínimo de erro que o algoritmo busca. Um algoritmo como o *Backpropagation* varia seu gradiente até ele convergir em uma região de erro mínimo, que pode ser global ou local. A Figura 3.10 relaciona um exemplo de gradiente convergindo para uma região de mínimo.

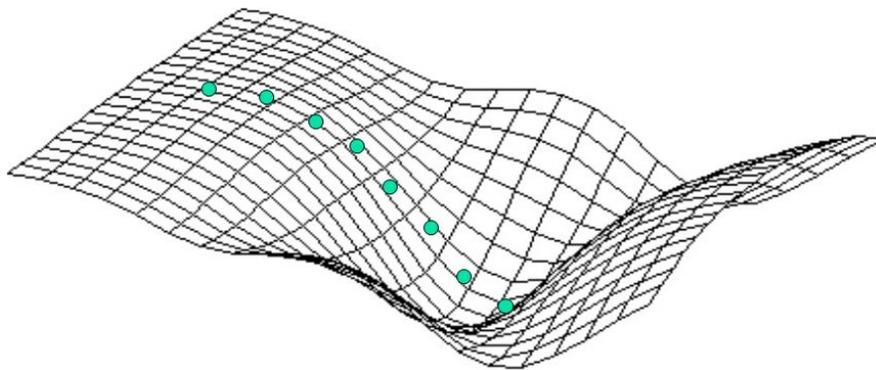


Figura 3.10: Exemplo de convergência do Vetor Gradiente

Os passos do *Backpropagation* são expostos em seu algoritmo [68] [70]. Após a

inicialização dos pesos da rede neural, inicia-se a chamada fase “*feedforward*”, que é a etapa do treinamento em que os dados de entrada são propagados através das sinapses e neurônios da rede. O valor  $h_j$  é dado pelo somatório da multiplicação dos valores que são propagados com o peso da sinapse que o transporta. Sobre este valor  $h_j$  (equação 3.1), aplica-se a função de ativação  $g(x)$  escolhida.

O processo é repetido por cada camada da rede, até que se atinja sua camada de saída. Neste momento, termina-se a fase de propagação “*feedforward*” e inicia-se a fase *backward*. Em tal fase, é feito o cálculo do erro gerado pela propagação da entrada, havendo diferentes medições de erro aplicadas no seu treinamento, como por exemplo, o erro médio quadrático ou a soma dos quadrados [58]. A função de erro é apresentada de forma genérica em nosso algoritmo, na equação 3.5.

Após calculado o erro, as equações 3.9 e 3.11 indicam a fase de atualização dos pesos das sinapses. O fator  $\eta$  também é parametrizável e define a intensidade das atualizações dos pesos. Um valor alto de fato pode apresentar um rápido treinamento, como também fazer com o treinamento não alcance o mínimo global. Um valor baixo para  $\eta$  faz com que o treino se torne mais lento, mas aumenta as possibilidades do treinamento convergir para um mínimo global. Por fim, repete-se todo o processo, propagando novamente os dados de entrada, em uma nova ordem selecionada aleatoriamente.

A seguir, são relacionadas as variáveis usadas no algoritmo e seus respectivos significados:

- $n$ : número de neurônios na camada escondida considerada.
- $m$ : número de entradas (considerando o *bias*) da rede.
- $L$ : número de saídas da rede.
- $v_{ij}$ : pesos sinápticos ligando a entrada  $i$  ao neurônio  $j$  da camada escondida.
- $x_i$ : pesos sinápticos ligando a entrada à camada escondida.
- $g(x)$ : função de ativação do neurônio.
- $w_{jk}$ : pesos sinápticos da camada anterior, ligando o neurônio  $j$  ao neurônio  $k$ .
- $E$ : Função de Erro entre a saída do neurônio da última camada e o valor alvo.
- $\Delta v_{ij}$ : Variação do peso  $v_{ij}$ .
- $\Delta w_{jk}$ : Variação do peso  $w_{jk}$ .

São expostas a seguir as fases “*forward*” e “*backward*” do *Backpropagation*.

---

**Algoritmo 1:** Algoritmo *Backpropagation - Forward*

---

**início**

Inicializa todos os pesos com valores randômicos;

**Treinamento**

**repita**

**Forward Fase**

Computa a ativação de cada neurônio  $j$  na camada escondida, considerando  $i = 0$  como sendo o *bias*, usando:

$$h_j = v_{0j} + \sum_{i=1}^m x_i v_{ij} \quad (3.1)$$

$$a_j = g(h_j) \quad (3.2)$$

Propaga através da rede até chegar na camada de saída, que possui os neurônios  $k$ . A entrada  $j = 0$  no neurônio  $k$  representa o *bias* da camada. As ativações são:

$$h_k = w_{0k} + \sum_{j=1}^n a_j w_{jk} \quad (3.3)$$

$$y_k = g(h_k) \quad (3.4)$$

**até treinamento;**

**fim**

---

---

**Algoritmo 2:** Algoritmo *Backpropagation* - Backward

---

início

**Treinamento**

**repita**

**Backward Fase**

Computa o erro da saída usando

$$E = f(t_k - y_k) \quad (3.5)$$

$$\delta_{ok} = E g'(y_k) \quad (3.6)$$

Computa o erro na camada escondida usando

$$\delta_{hj} = g'(a_j) \sum_{k=0}^L w_{jk} \delta_{ok} \quad (3.7)$$

Atualiza os pesos da camada de saída usando

$$\Delta w_{jk} = \eta \delta_{ok} a_j^{hidden} \quad (3.8)$$

$$w_{jk} \leftarrow w_{jk} + \Delta w_{jk} \quad (3.9)$$

Atualiza os pesos da camada escondida usando

$$\Delta v_{ij} = \eta \delta_{hj} x_i \quad (3.10)$$

$$v_{ij} \leftarrow v_{ij} + \Delta v_{ij} \quad (3.11)$$

**até** *treinamento terminar*;

**Retorna a fase *Forward***

fim

---

### 3.1.6 Algoritmo *Resilient Backpropagation*

O algoritmo *Resilient Backpropagation* foi desenvolvido com o objetivo de minimizar o tempo de aprendizagem do *Backpropagation*. Seu objetivo é eliminar o efeito causado pelo valor das derivadas parciais, usadas na determinação do vetor gradiente no *Backpropagation* [68].

O tamanho da atualização de um peso sináptico depende não só da taxa de aprendizado, mas também da derivada parcial, assim o valor imprevisível desta pode afetar na escolha da taxa de aprendizado. O *Resilient Backpropagation* propõe

então a atualização do peso sem considerar o valor da derivada parcial mas sim apenas o seu sinal, indicando a direção do gradiente [69].

Ao invés de utilizar o valor da derivada parcial, cada peso tem um valor individual, denotado por  $\Delta_{ij}$ , sendo esse valor modificado durante a aprendizagem e baseado no sinal da derivada parcial correspondente à variação do valor do peso em questão.

Considerando primeiramente esse valor individual de atualização, ele é baseado nos instantes de tempo do treinamento (épocas) e na função de erro  $E = f(t_k - y_k)$  definida para o treinamento. A taxa de aprendizado  $\eta$  é definida no *Resilient Backpropagation* como  $\eta^+$  e  $\eta^-$ , levando em consideração a direção do vetor gradiente. Os valores de aprendizado seguem a condição :

$$0 < \eta^- < 1 < \eta^+$$

A atualização acontece seguindo a seguinte regra: se o valor da derivada se mantém por duas iterações consecutivas, o valor individual de atualização do peso é atualizado pelo fator  $\eta^+$ , indicando que não houve mudança no sinal do gradiente. Já quando a derivada muda de sinal, o valor da atualização é feita pelo fator  $\eta^-$  [69] e isso indica que o gradiente acabou de passar por uma região de mínimo local. Caso não haja mudança no gradiente, o valor de  $\Delta_{ij}$  não é alterado. Essa regra é apresentada na equação 3.12, abaixo:

$$\Delta_{ji}(t) = \begin{cases} \eta^+ * \Delta_{ij}^{(t-1)}, & \text{se } \frac{\partial x}{\partial w_{ij}}^{(t-1)} * \frac{\partial x}{\partial w_{ij}}^{(t)} > 0 \\ \eta^- * \Delta_{ij}^{(t-1)}, & \text{se } \frac{\partial x}{\partial w_{ij}}^{(t-1)} * \frac{\partial x}{\partial w_{ij}}^{(t)} < 0 \\ \Delta_{ij}^{(t-1)}, & \text{caso contrário} \end{cases} \quad (3.12)$$

Uma vez que o valor para a atualização dos pesos está definido, a definição do sinal que ele assume é feita seguindo uma regra simples, onde caso a derivada seja positiva (aumento do erro) os pesos  $w_{ij}$  tem o valor de  $\Delta_{ij}$  subtraídos, enquanto que no caso da derivada ser negativa, o valor de  $\Delta_{ij}$  é somado ao  $w_{ij}$ . Isso mostra que a atualização dos pesos segue a direção contrária ao vetor gradiente. Se não há alteração nas derivadas parciais, o valor definido é 0. A equação 3.13 relaciona a regra de de determinação do sinal de  $\Delta_{ij}$ .

$$\Delta w_{ij}^{(t)} = \begin{cases} -\Delta_{ij}^{(t)}, & \text{se } \frac{\partial x}{\partial w_{ij}}^{(t)} > 0 \\ +\Delta_{ij}^{(t)}, & \text{se } \frac{\partial x}{\partial w_{ij}}^{(t)} < 0 \\ 0, & \text{caso contrário} \end{cases} \quad (3.13)$$

No entanto, há uma exceção para a definição do valor de  $\Delta w_{ij}^{(t)}$ , especificamente quando há uma mudança de sinal na derivada parcial. Esta mudança indica que o

“passo” (ou a atualização) foi muito grande, fazendo com que o valor do mínimo global ou local não fosse alcançado. Assim, o antigo valor de  $\Delta w_{ij}^{(t)}$  é revertido e redefinido pela equação 3.14, abaixo:

$$\Delta w_{ij}^{(t)} = -\Delta w_{ij}^{(t-1)}, \frac{\partial x^{(t)}}{\partial w_{ij}} * \frac{\partial x^{(t)}}{\partial w_{ij}} < 0 \quad (3.14)$$

Por fim, é feita a atualização dos pesos é feita efetivamente pela equação 3.15, atualizando seus valores para a próxima época de treinamento.

$$w_{ij}^{(t+1)} = w_{ij}^{(t)} + \Delta w_{ij}^{(t)} \quad (3.15)$$

Portanto, a lógica do algoritmo *Resilient Backpropagation* 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 [69].

### 3.1.7 Redes de Operação

As Redes Neurais são ferramentas poderosas e sua capacidade de resolução de problemas complexos faz com elas tenham as mais diversas aplicações. Elas são usadas para diferentes estudos, como também são usadas de maneira definitiva como uma ferramenta de apoio ao problema. Sendo assim, apresenta-se o conceito de “Redes de Operação”.

As “Redes de Operação” são importantes no cenário das RNAs, sendo aplicadas basicamente quando quer se utilizar a rede neural (ou sistema inteligente) já treinada para avaliar amostras reais e atuar sobre o problema. Um exemplo deste conceito pode ser visto no uso de RNAs usadas como ferramenta de apoio ao diagnóstico de pacientes com Tuberculose (*M. tuberculosis*) [61] ou aplicados no cenário de recomendação de filmes a um cliente [62] por uma empresa do setor. Em ambos os exemplos, um modelo específico foi escolhido para atuar sobre o problema em um cenário real, e portanto o que temos é que uma Rede Operação é a RNA que é escolhida para ser usada em “campo”.

Em situações práticas, são geradas diferentes RNAs para um mesmo problema, alterando-se por exemplo a quantidade de neurônios em suas camadas, as funções de ativação, ou mesmo a distribuição dos dados em seus grupos de treinamento, validação e teste. Cada problema possui ainda características que influenciam na escolha desta rede, como por exemplo se seus dados são representativos ao espaço de amostras.

Sendo assim, a escolha de uma única rede neural para ser posta em operação busca definir a rede que melhor atende aos seus objetivos, que variam de acordo

com o problema.

## 3.2 Métodos Estatísticos

A avaliação de um método de Aprendizado de Máquina pode ser feita de diferentes formas e uma das abordagens para esta avaliação é a abordagem estatística. Nesta seção são apresentados alguns dos métodos estatísticos usados na análise de sistemas de Aprendizado de Máquina e que também foram utilizados neste estudo.

São expostos nas subseções a seguir as chamadas medidas de Sensibilidade, Especificidade e Acuidade. Os métodos seguintes são a chamada curva ROC e a medida SP (Soma-Produto). Por último, apresenta-se a técnica de Validação Cruzada.

### 3.2.1 Acurácia, Especificidade e Sensibilidade

As medidas de Sensibilidade e Especificidade são conceitos largamente utilizados na área médica, principalmente para o diagnóstico e identificação de uma certa condição médica, que neste cenário pode possuir apenas dois estados possíveis: Presente ou Ausente.

Uma aplicação comum desses conceitos é feita em exames médicos para algum diagnóstico, onde caso a condição testada esteja presente o resultado é considerado “Positivo”, caso contrário ele é considerado como “Negativo”. No entanto, esses dois resultados geram quatro possibilidades distintas para a observação e identificação da condição testada. A condição observada se refere ao resultado que um exame retorna à condição médica, enquanto a real é o resultado que realmente representa o estado da amostra. A Figura 3.11 apresenta graficamente estas possibilidades, enquanto a Tabela 3.1 relaciona estes valores de acordo com sua condição.

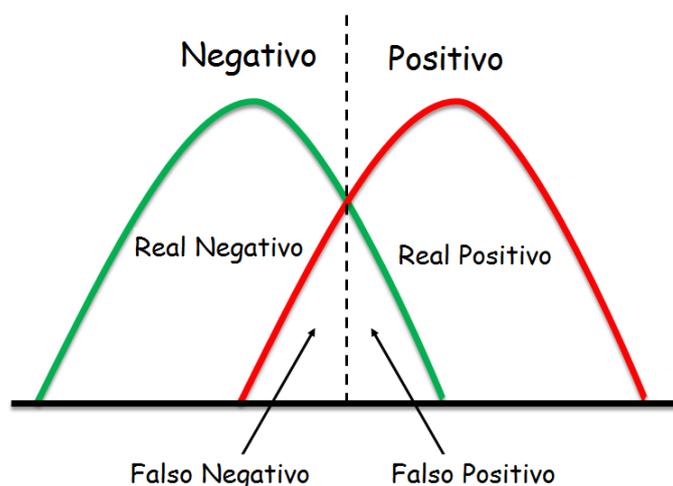


Figura 3.11: Resultados possíveis para as condições Negativa e Positiva

Tabela 3.1: Possíveis resultados entre a observação e a real condição

		Condição Real	
		Positivo	Negativo
Condição Observada	Positivo	Real Positivo (RP)	Falso Positivo (FP)
	Negativo	Falso Negativo (FN)	Real Negativo (RN)

Uma observação é chamada de “Real Positivo” (RP) quando o resultado encontrado e a real condição são positivas, enquanto que a chamada “Falso Negativo” (FP) é assim definida quando o resultado encontrado é positivo mas seu real estado é negativo. Uma observação “Falso Negativo” (FN) é vista quando a condição observada é negativa e sua condição real é positiva. Um “Real Negativo” (RN) é dado quando tanto a observação quanto a o estado real são negativos.

A Sensibilidade é a medida estatística que relaciona apenas a condição “Positivo” das amostras. A medida retorna a proporção de amostras que são corretamente identificados. No cenário abordado, esta medida pode representar a proporção de amostras positivas que um exame médico consegue identificar corretamente. O seu cálculo é feito dividindo a quantidade de resultados “Real Positivo” sobre a quantidade total dos resultados considerados positivos, como na fórmula abaixo:

$$\text{Sensibilidade} = \frac{RP}{(RP + FN)} \quad (3.16)$$

Já a Especificidade é a medida que relaciona apenas a condição “Negativo” das amostras. A medida retorna a proporção entre as amostras corretamente identificadas com todas as amostras classificadas como negativa, certas ou erradas. Seu cálculo é feito dividindo a quantidade de resultados “Real Negativo” sobre a quantidade total dos resultados considerados negativos, como na fórmula abaixo:

$$\text{Especificidade} = \frac{RN}{(RN + FP)} \quad (3.17)$$

A medida de Acurácia é mais abrangente que a Sensibilidade e a Especificidade, englobando as duas condições consideradas: “Positivo” e “Negativo”. Ela retorna a proporção de amostras que foram identificadas corretamente, entre todas as que foram testadas, para ambas as condições. Um exemplo de aplicação da medida seria para verificar a proporção de pacientes que um exame diagnostica corretamente, independente de seu real estado. Para calculá-la deve-se somar a quantidade de resultados “Real Positivo” e “Real Negativo” e dividi-los pela quantidade total de resultados. A fórmula que representa esse conceito é descrita abaixo.

$$\text{Acurácia} = \frac{RP + RN}{(RP + RN + FN + FP)} \quad (3.18)$$

Os conceitos de Sensibilidade, Especificidade e Acurácia são usados também na área de Aprendizado de Máquina, principalmente em problemas de classificação binária. O seu uso em outras áreas gera apenas uma mudança na terminologia, onde por exemplo, a Sensibilidade é referida como “Taxa de Verdadeiro Positivo” ou “Revocação” (*Recall*), a Especificidade como “Taxa de Verdadeiro Negativo” ou “Revocação Inversa” (*Inverse Recall*), e a Acurácia como “Acuidade”. Outros termos utilizados são: “Acerto”, “Erro” e “Falso Alarme” para respectivamente, “Real Positivo”, “Falso Negativo” e “Falso Positivo” [71].

### 3.2.2 Curva ROC

A curva ROC teve sua origem durante a Segunda Guerra para analisar sinais eletrônicos e problemas com radares. Ela era utilizada para medir quanto um operador sabia distinguir entre ruídos vindos de um avião inimigo e um ruído aleatório, por exemplo. Sua sigla tem origem nesse contexto e significa *Receiver Operating Characteristic* [72].

A curva ROC é uma técnica e ferramenta de análise com aplicação em diferentes problemas e áreas do conhecimento. Seu uso é feito por exemplo, em problemas de detecção de sinais e de falso alarme, na área de Aprendizado de Máquina, especialmente em problemas de classificação e tomada de decisão binária ou com mais classes [73]. Outro caso em que essas curvas são utilizadas são problemas na área da medicina, com aplicação para testes de diagnósticos de diferentes enfermidades [74] [75].

O gráfico de uma ROC é bidimensional, com seu eixo das abscissas definido para representar valores de Sensibilidade (Real Positivo) e seu eixo das ordenadas representando valores de Falsos Positivos ( $1 - \text{Especificidade}$ ). O objetivo da curva é relacionar os ganhos e perdas entre a Sensibilidade e a Especificidade, e por causa destas é delimitado que seus dois eixos possuam seus valores variando de 0 até 1 apenas. A Figura 3.12 apresenta exemplos de curva ROC.

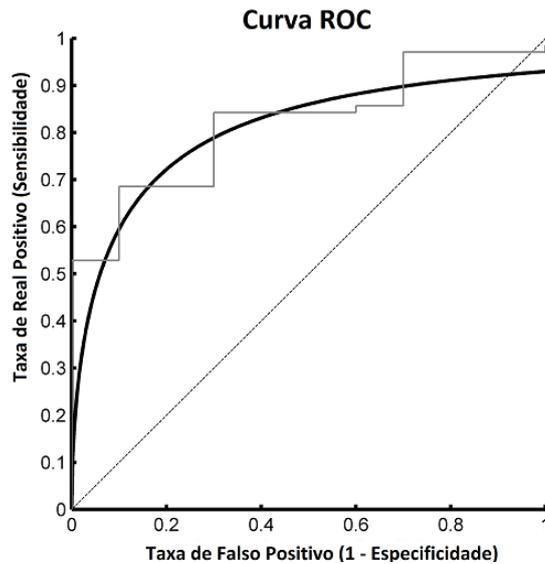


Figura 3.12: Exemplos de curva ROC

A figura apresenta duas curvas ROC, sendo a primeira curva em um formato contínuo e a segunda em um formato discreto. Ambas são geradas por um conjunto de valores de limiares ou separadores que dividem os resultados (ou classes) em por exemplo, negativo ou positivo, doente ou sadio, classe A ou classe B, dependendo do problema analisado pela curva. Cada limiar tem associado a ele um valor de Sensibilidade e Especificidade, que gera os gráficos vistos.

A diferença no formato entre as curvas é dado pela quantidade de valores de limiar que foram testados. A curva em formato discreto é gerada por uma quantidade limitada de valores de separador (número finito de pontos) e esta condição implica que qualquer curva ROC apresentará o formato de uma função degrau [73]. Esses casos representam curvas ROC reais, onde sempre haverá uma quantidade limitada de dados. Já a curva contínua representa a função ou comportamento que a curva discreta aproxima, e conceitualmente esta curva poderia ser gerada apenas se os valores testados fossem os componentes do intervalo  $(-\infty, +\infty)$ . Logo, esta segunda curva não é vista em cenários práticos.

Outro fator importante para o uso da curva ROC se refere a análise que é feita sobre ela. A Figura 3.13 apresenta o espaço onde a curva está inserida e apresenta diferentes pontos neste espaço e suas interpretações.

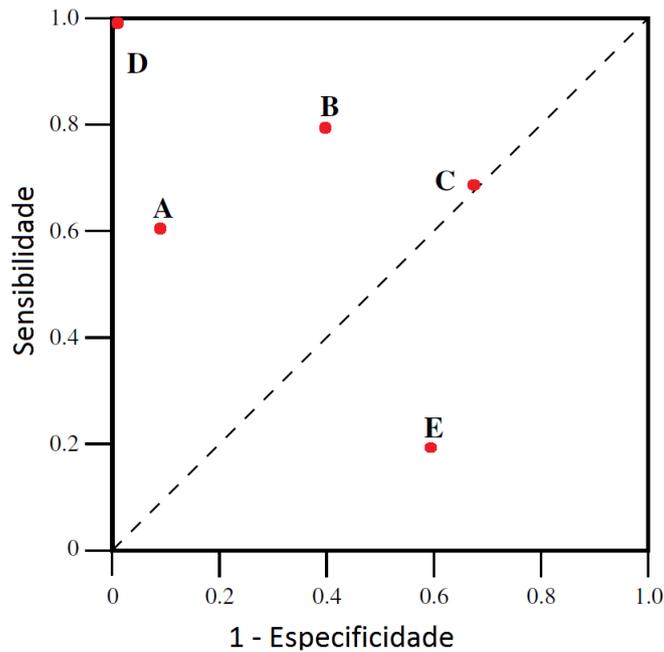


Figura 3.13: Exemplos de pontos no espaço de uma curva ROC

O ponto **D** (0, 1) é o ponto que caso uma curva ROC passe por ele, ela é considerada como a curva perfeita, com o valor máximo de Sensibilidade e Especificidade. O ponto **C** (0.7, 0.7), sobre a reta que liga os pontos (0, 0) e (1, 1) (reta  $y = x$ ) indica que o valor de Sensibilidade e Falso Alarme são idênticos. Esse ponto em casos de classificadores binários por exemplo, indica que a tomada de decisão entre duas classes é randômica em 70% dos casos. O ponto **E** na Figura 3.13 é apresentado na região triangular abaixo da reta anteriormente citada, e este ponto indica que uma curva que por ele passe apresenta um poder de decisão pior que uma decisão randômica. O que se vê na prática é que este ponto inverte por exemplo, as classificações. Um exemplo seria o ponto (1, 0), onde caso um classificador binário apresentasse tal resultado, ele sistematicamente classificaria as amostras com a classe errada [73]. No entanto, tais casos são incomuns em análises de ROC.

Os pontos **A** (0.1, 0.6) e **B** (0.4, 0.8) são exemplos de pontos mais comumente encontrados em análises por curvas ROC. A maioria dos pontos de uma ROC se encontram na região do triangulo superior do espaço da curva, e há diferentes interpretações para eles. O ponto **A** é um ponto que apresenta uma Sensibilidade menor do que **B**, mas também apresenta uma menor taxa de Falso Alarme em comparação com o segundo ponto. A definição de qual ponto possui uma melhor performance depende do problema analisado, uma vez que as características de cada um determina seu objetivo, podendo ser a maximização de sua Sensibilidade ou minimização de sua taxa de Falso Alarme.

Além dos pontos onde passam as curvas, outras medidas são consideradas para análise e principalmente comparação entre curvas ROC. Uma medida usada é a

chamada **AUC** (*Area Under the Curve*) ou a “Área Abaixo da Curva”. A medida quantifica a área que está abaixo de uma curva ROC, e quanto maior o seu valor melhor é o desempenho de uma curva em comparação a outra. Uma vez que os valores nos eixos do gráfico da ROC variam 0 até 1, a área máxima deve ser igual a 1. Ao considerar a reta ( $y = x$ ), ela divide o espaço da ROC em duas áreas idênticas, então neste caso sua área é igual a 0.5. Métodos de decisão totalmente aleatórios apresentam este valor para a medida AUC. Os casos mais comuns de ROCs são casos em que suas medidas de AUC estão entre os valores de 0.5 e 1. A Figura 3.14 trás exemplos de curvas ROC e suas medidas AUC para os três casos descritos.

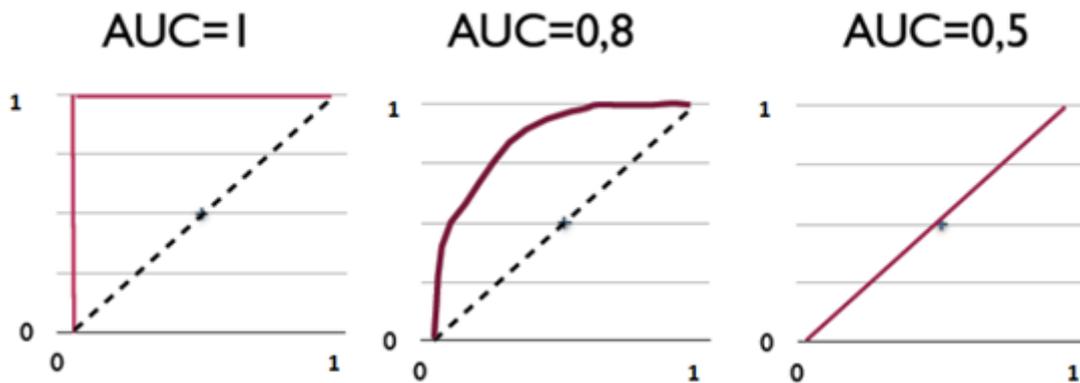


Figura 3.14: Exemplos de medidas **AUC** em curvas ROC

O cálculo de AUC pode ser feito de diferentes maneiras e possui relação com a estatística U de Mann-Whitney ou dos resultados do teste de Wilcoxon [76] [75]. O seu valor pode ser calculado ainda a partir da medida de dispersão estatística de coeficiente de Gini [73]. No entanto, o cálculo mais comumente feito para AUC é usando uma regra de integração sobre a curva ROC, como por exemplo pela regra do trapézio [60]. Para esta técnica de integração, a curva é interpolada para um conjunto de pontos. Em situações práticas estes pontos são os valores usados para gerar a curva, formando-se trapézios abaixo deles. Cada trapézio tem sua área calculada e ao fim, a soma de todos estes valores aproximam o valor da área e retornam a medida AUC. Na Figura 3.15 é visto um exemplo de curva ROC e os trapézios destacados para seus pontos interpolados, e para este exemplo a área abaixo da curva é dada por ( $AUC = A1 + A2 + A3 + A4$ ).

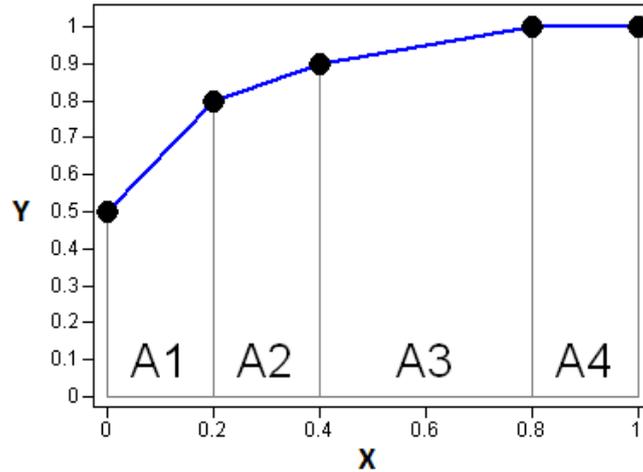


Figura 3.15: Exemplo de Integração pela Regra do Trapézio

A análise dos valores de AUC são intrinsecamente dependentes do problema, uma vez que o mesmo resultado pode ser considerado bom ou ruim dependendo do contexto. Logo não há um valor padrão para a avaliação de performance por esta medida.

### 3.2.3 Valor SP

Uma outra medida de avaliação de performance é o chamado *Valor SP* ou *Valor Soma-Produto*, e tal medida é uma alternativa ao uso de curvas ROC. O primeiro fator que diferencia o SP da análise por ROC é que esta medida é um valor numérico apenas, e quantifica a relação entre as taxas de Sensibilidade (S) e a Especificidade (E). O cálculo do valor SP é dado por [77]:

$$SP = \sqrt{\left(\frac{S + E}{2}\right) \times \sqrt{S \times E}} \quad (3.19)$$

A partir da equação 3.19 é possível observar características sobre o comportamento do SP. O primeiro ponto visto é que, uma vez que a Sensibilidade e a Especificidade variam no intervalo de 0 até 1, o valor SP também varia dentro dessa faixa dinâmica. A Figura 3.16 mostra a variação do valor SP pelos valores de Sensibilidade e Especificidade.

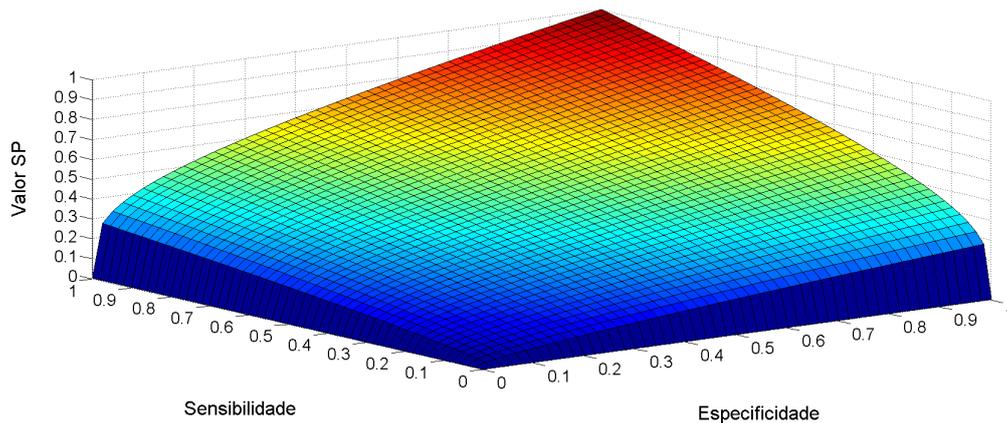
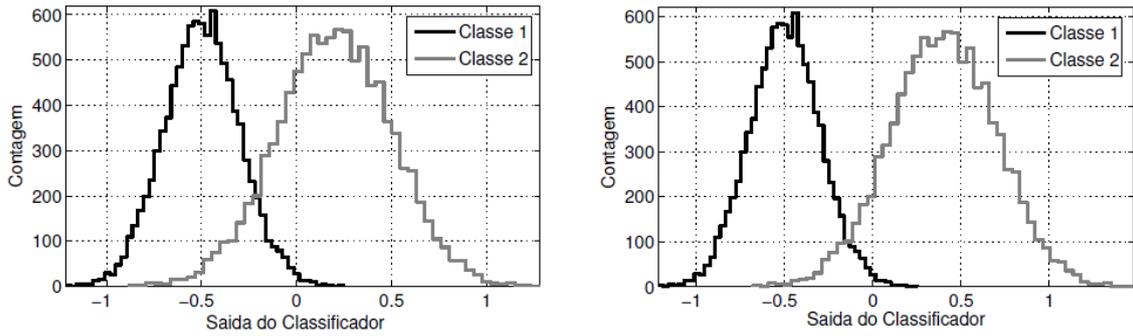


Figura 3.16: Variação do Valor SP baseado nos valores da Sensibilidade e da Especificidade

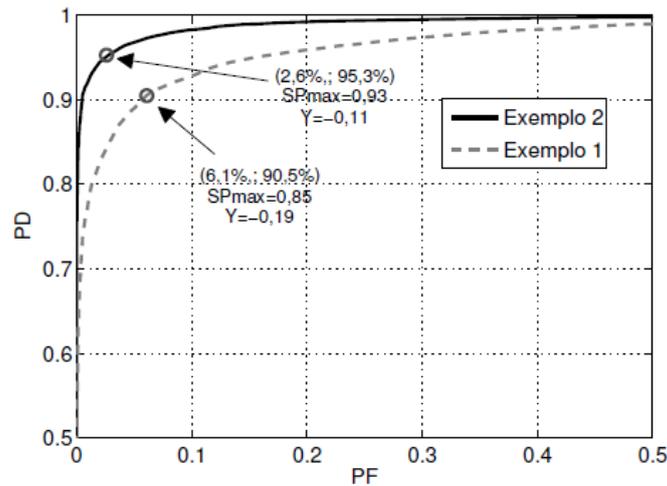
O comportamento visto na figura mostra primeiramente que o SP busca equilibrar a Sensibilidade e a Especificidade, sendo isto percebido quando as medidas apresentam valores distantes entre si. Um exemplo para esta situação é quando uma das medidas apresenta seu valor máximo (1) e a outra com seu valor mínimo (0), retornando um valor SP também igual a 0. A medida que os valores aumentam e se mantêm próximos, o valor SP também aumenta, atingindo seu valor máximo quando ambas as medidas também o apresentam. A medida SP ainda assume que tanto a Sensibilidade e a Especificidade possuem a mesma importância ou peso para seu cálculo, o que reforça o conceito de equilíbrio entre as medidas e indica que seu uso não é ideal para problemas de maximização de taxas de Real Positivo ou minimização de taxas de Real Negativo apenas.

O valor SP é empregado principalmente em problemas de classificação binária ou de discriminadores, onde há confusão na definição ou separação de duas classes. Em tais situações, configura-se os casos com *trade-off*, ou seja, em que necessariamente o ganho de uma classe representa a perda da outra. Nesses casos, o objetivo é de minimizar a perda em ambas classes.

O SP pode ser usado ainda como medida de comparação entre classificadores, como exemplificado por FILHO [2], onde foi apresentado um cenário hipotético com dois classificadores e em seguida foram obtidas suas medidas SP. A Figura 3.17 relaciona a distribuição das saídas dos dois classificadores hipotéticos e de sua análise por valor SP e curva ROC.



(a) Distribuição das saídas dos Classificadores 1 (à Esquerda) e 2 (à Direita)



(b) Curvas ROC e Valores SP dos classificadores

Figura 3.17: Distribuição e Resultados do cenário hipotético com análise por ROC e SP [2]

Na Figura 3.17(a), os dois gráficos possuem uma região de confusão entre suas saídas, indicando que amostras de ambas as classes possuem semelhanças para o classificador. No entanto se percebe que o classificador 1, à esquerda, possui uma região de confusão menor que o segundo classificador, à direita. Isso implica em que o primeiro possua uma melhor capacidade de diagnóstico correto comparado ao segundo. Na Figura 3.17(b) são relacionadas as curvas ROC de cada um deles e a medida SP com melhor performance, que representa a melhor separação possível para eles. O valor SP do primeiro igual 0,93 contra 0,85 do segundo, conseguindo refletir a melhor taxa de Real Positivo deste classificador. Outro fator é a concordância entre o valor SP e a medida AUC (Área Abaixo da Curva) das ROCs, onde também o valor AUC do classificador 1 é maior que a do classificador 2, mostrando que a análise por SP possui a mesma capacidade a análise por curvas ROC.

### 3.2.4 Validação Cruzada

A Validação Cruzada é uma técnica usada para avaliar o erro estatístico de qualquer sistema de decisão e também é considerada uma técnica de validação de modelos. Um de seus objetivos é avaliar a precisão de um modelo preditivo, e seu conceito está relacionado com o conjunto de dados usado pelo sistema inteligente assim como aos conceitos de grupos de Treinamento, Validação e Teste, vistos na Seção 3.1.

A escolha dos conjuntos de dados para a fase de treinamento e aprendizado é um fator importante para um sistema inteligente, e a Validação Cruzada é aplicada de modo a definir como um conjunto de dados será dividido e utilizado tanto no treinamento, como na estimação do erro do sistema de decisão. Três diferentes métodos de Validação são largamente conhecidos: o método *Holdout*, o método de *k-Fold* e o método *Leave-one-out*. Para a exposição dos métodos a seguir, considere que o chamado “conjunto de teste” é formado pelos grupos de dados de validação e teste.

O método *Holdout* é o mais simples dentre os três mencionados e consiste em dividir, uma vez apenas, todo o conjunto de dados disponível em dois grupos mutuamente exclusivos: Conjunto de treinamento e Conjunto de teste. A proporção de dados alocados para cada grupo ou mesmo a quantidade de amostras que cada um possui depende de fatores como por exemplo, o problema estudado, a quantidade de dados disponíveis e se os dados estão balanceados entre suas classes ou definições. Com os dados divididos, o conjunto de treinamento é utilizado exclusivamente para o aprendizado da rede, enquanto o outro conjunto é usado para avaliar sua performance. Seu uso é recomendado quando há disponível um grande conjunto de dados para o problema [78]. A Figura 3.18 esquematiza a divisão dos dados no método *Holdout*.



Figura 3.18: Método *Holdout* de Validação Cruzada

Os dados que são apresentados ao sistema, tanto para o seu aprendizado como

para sua validação, são importantes e podem gerar impactos na performance dos modelos [58]. Uma limitação do método é o fato de fazer apenas uma divisão em seus dados, o que pode fazer com um dos conjuntos (Treinamento ou Teste) apresente mais classes que o outro grupo, ou que os dados alocados em um dos grupos sejam dados que não representem todo o conjunto original. Tal fato acarreta em um mau aprendizado do sistema inteligente e conseqüentemente, um maior erro em suas saídas.

O método *K-fold* busca superar a limitação vista no *Holdout*, com sua única divisão do conjunto de dados. O método consiste em dividir os dados randomicamente em  $k$  subconjuntos (ou *folds*) distintos e de mesmo tamanho, e fazer diferentes treinamentos com diferentes combinações dos subconjuntos considerados. Entre os  $k$  subconjuntos gerados,  $k - 1$  são selecionados e definidos para compor o grupo de treinamento, e o restante é utilizado como conjunto de teste. O processo é repetido  $k$  vezes, onde a cada etapa um novo subconjunto é definido para o conjunto de teste. Em cada repetição é treinado um sistema inteligente e são calculadas suas performances, e ao fim do processo consegue-se estimar, por exemplo, média e desvio padrão da performance de todo o modelo. A Figura 3.19 esquematiza e exemplifica a divisão feita no método *K-fold*. No exemplo visto na figura, o conjunto de dados foi dividido em 5 subconjuntos e sendo assim, o conjunto de treinamento foi formado por 4 *folds* e o conjunto de teste com o restante.

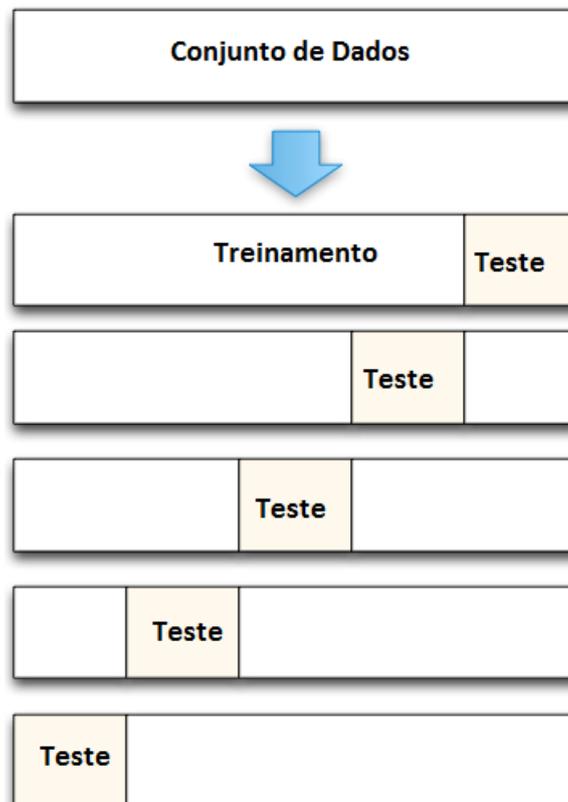


Figura 3.19: Método *K-fold* de Validação Cruzada

O principal custo dessa técnica é computacional, gerado pelo maior número de treinamentos aos quais são submetidos os dados. Nesse método são feitos  $k$  treinamentos diferentes, enquanto que no método *Holdout* apenas um treinamento é necessário.

O método *Leave-one-out* é um caso específico do método *K-fold* e sua diferença está na atribuição e determinação dos *folds*. Esse método consiste em definir como número de subconjuntos o valor igual à quantidade de dados disponíveis, ou seja, cada subconjunto é formado por apenas uma amostra que compõe os dados. Após a definição dos  $k$  subconjuntos, o processo é o mesmo definido para o *K-fold*, escolhendo-se  $k - 1$  subconjuntos para treinamento e deixando o último como conjunto de teste. Tomando como exemplo a situação ilustrada na Figura 3.19, no método *leave-one-out* o conjunto de dados apresentaria um total de 5 amostras e seria definido que o conjunto de treinamento seria formado por 4 amostras de dados e a última seria usada para teste e validação. O custo computacional deste método aumenta, se comparado ao método *K-fold*, sendo seu uso indicado para situações onde poucos dados estão disponíveis [78].

# Capítulo 4

## Método Experimental

Neste capítulo são explicadas as abordagens, detalhados os conjuntos de dados e os experimentos conduzidos neste estudo, buscando responder aos objetivos estabelecidos na Seção 1.2. O primeiro experimento busca analisar qual o impacto causado pela adição de novos exames de identificação da tuberculose aos hoje usados. Em seguida, é feita uma segunda abordagem para avaliar o efeito do primeiro experimento. No último experimento, é feito um estudo baseado apenas nos exames preconizados pela legislação brasileira e seu objetivo é buscar otimizá-los e aumentar suas eficiências.

A Seção 4.1 deste capítulo descreve os conjuntos de dados utilizados no estudo. As seções 4.2 e 4.3 descrevem os experimentos feitos para a análise do acréscimo de novos exames na identificação de tuberculose. A Seção 4.4 descreve o experimento de otimização dos testes cervicais.

### 4.1 Dados

A seção expõe uma descrição dos conjuntos de dados disponibilizados para este estudo, como também aborda o formato padrão que os dados irão assumir. Foram fornecidas três bases de dados, provenientes de diferentes fazendas. O primeiro conjunto de dados corresponde a animais localizados em uma fazenda de propriedade da Universidade Federal Fluminense - UFF. O segundo conjunto é pertencente a uma fazenda localizada no município de Macaé, estado do Rio de Janeiro, sendo ela uma propriedade privada. O terceiro é relativo a um conjunto de propriedades particulares localizadas no estado do Rio Grande do Sul, e foi fornecido pela EMBRAPA.

Os conjuntos de dados são compostos por valores numéricos e categóricos. A Tabela 4.1 relaciona os valores nominais presentes nas bases de dados e seus respectivos significados. Os valores **Pos** e **Neg** indicam que o exame feito no animal apresentou, respectivamente, um diagnóstico positivo e negativo para tuberculose bovina. O

valor **Inc** é apresentado quando o resultado do exame é considerado inconclusivo. O valor **Susp** é usado nas situações em que há a suspeita de que o animal seja anérgico. Por fim, o valor ‘-’, que significa Não se Aplica, é usado quando não é necessária a realização de tal exame no animal.

Tabela 4.1: Relação de valores e significados dos campos dos conjuntos de dados

<b>Representação</b>	<b>Significado</b>
<b>Pos</b>	Positivo
<b>Neg</b>	Negativo
<b>Inc</b>	Inconclusivo
<b>Susp</b>	Suspeito
-	Não se aplica

As subseções a seguir descrevem cada uma das bases de dados e as adaptações feita sobre elas, para serem usadas pelos modelos neurais.

#### 4.1.1 Fazenda UFF

A Fazenda UFF refere-se à Fazenda Escola de Cachoeiras de Macacu (FECM) da Universidade Federal Fluminense, sendo esta unidade responsável por atividades de pesquisa, ensino e extensão, além da produção de alimentos de origem animal. Para o estudo, tal unidade pode ser considerada uma fazenda modelo por se tratar de um ambiente controlado, com maior rigor e precisão na execução dos testes e das análises. Seus dados são datados de Abril de 2013.

A Tabela 4.2 traz a relação de animais por diagnóstico na base de dados. A fazenda possui um total de 77 animais, com diagnósticos definidos para infecção, ou não, de Tuberculose Bovina. Os animais diagnosticados com tuberculose representam aproximadamente 40% do total de animais (31 animais), enquanto os animais diagnosticados como sadios e livres de tuberculose contabilizam 46, representando 60% do total. Os animais anérgicos estão incluídos no grupo de animais com tuberculose bovina.

Tabela 4.2: Quantidade de Animais TB+ e TB− para a fazenda Fazenda UFF

<b>Status TB</b>	<b>Ocorrências</b>	<b>Porcentagem</b>
Infectado	31	40%
Não Infectado	46	60%
<b>Total</b>	<b>77</b>	<b>100%</b>

Cada amostra do conjunto de dados representa um animal analisado na FECM e

cada uma possui um conjunto de exames para o referido animal, com seus resultados ou diagnósticos. Os dados que compõem o conjunto podem possuir valores numéricos ou categóricos. A Tabela 4.3 apresenta relação dos testes disponíveis para a fazenda, o formato do campo e os valores que cada um pode assumir. Para as variáveis do tipo numérico, é apresentado o intervalo de variação de seus valores.

A base possui valores numéricos com as medidas dos animais aplicados com PPD Bovina e PPD Aviária. Os valores apresentados representam a variação do volume da região em que foram aplicadas as tuberculinas, medidas em milímetros. Tais dados permitem a definição de diagnóstico para o TCS e o TCC dos animais, segundo a norma oficial [3].

Para o exame do tipo *ELISA*, a fazenda UFF possui três variedades: ELISA MPB70, ELISA MPB83 e ELISA ESAT-6. Cada um dos três ELISAs citados possui o seu valor de densidade ótica associado ao animal e também a indicação de quais destas saídas estão acima do ponto de corte, indicando resultado positivo para tuberculose. Os valores de densidade ótica dos recombinantes MPB70 e MPB83 variam de 0,05 até 3,50. Para o recombinante ESAT-6, a variação de valores de densidade ótica ocorre de 0,05 até 1,23.

A base dispõe ainda de 3 testes diferentes para os tecidos e leite dos animais. Para ambos (tecido e leite) são aplicados os testes de *Ziehl Neelsen* para cultura, a PCR direta sobre a amostra e por fim, a PCR sobre a cultura da amostra. Tais testes retornam valores positivos, negativos e inconclusivos para os exames, além de sua execução não ser necessária em alguns casos. Importante ressaltar que nem todas as amostras da base possuem o exame sobre o leite feito.

Tabela 4.3: Relação de exames usados, tipo do dado e seus possíveis valores - UFF

<b>Campo</b>	<b>Tipo</b>	<b>Valores</b>
<i>PPD Bov</i>	Numérico	[-4, 20 ; 13, 90]
<i>PPD Av</i>	Numérico	[-2, 60 ; 6]
ELISA MPB70	Numérico	[0, 05 ; 3, 50]
ELISA MPB83	Numérico	[0, 05 ; 3, 50]
ELISA ESAT-6	Numérico	[0, 05 ; 1, 23]
Cultura Leite	Categórico	Pos; Neg; Inc; '-'
PCR Cultura Leite	Categórico	Pos; Neg; '-'
PCR direta Leite	Categórico	Pos; Neg; '-'
Cultura Tecido	Categórico	Pos; Neg; Inc; '-'
PCR Cultura Tecido	Categórico	Pos; Neg; '-'
PCR direta Tecido	Categórico	Pos; Neg; '-'
Anérgicos	Categórico	Susp; Neg

Dentre os exames disponíveis para a fazenda UFF, três deles são considerados

como *padrão-ouro*: Cultura de tecido, PCR direta sobre as amostras de tecido e a PCR da cultura de tecidos. Esses exames são do tipo *post-mortem* e por isso, animais que apresentaram resultados negativos para os testes cervicais e para os testes ELISA não foram submetidos aos exames de tecido.

A base da UFF ainda apresenta informação de quais animais são suspeitos de serem anérgicos, sendo listado um total de 7 animais suspeitos. Todos os animais identificados, tiveram a presença de *M. bovis* confirmada através de seus exames de tecido. Na fazenda UFF, a prevalência de tuberculose é de 40% dos animais.

#### 4.1.2 Fazenda Macaé

A segunda fazenda analisada e descrita é uma propriedade privada, localizada no município de Macaé, estado do Rio de Janeiro. A Fazenda possui um total de 50 animais e seus diagnósticos definidos para infecção, ou não, de Tuberculose Bovina. A Tabela 4.4 traz a relação de animais por diagnóstico na base de dados. A distribuição de animais infectados e sadios é quase simétrica para esta base de dados. Do total de 50 animais, 23 (46%) foram diagnosticados com Tuberculose Bovina e os 27 animais restantes (54%) foram considerados livres da doença. Os animais suspeitos de anergia, estão incluídos no grupo de animais com tuberculose bovina.

Tabela 4.4: Quantidade de Animais TB+ e TB− para a fazenda Fazenda Macaé

Status TB	Ocorrências	Porcentagem
Infectado	23	46%
Não Infectado	27	54%
<b>Total</b>	50	100%

A Tabela 4.5 apresenta relação das informações disponibilizadas pela base de dados, tal como o tipo que cada campo assume e seus possíveis valores. Assim como a fazenda UFF, a fazenda Macaé possui os valores da PPD Bovina e PPD Aviária, podendo novamente se calcular as respostas dos animais para o TCS e o TCC. Seus valores representam a variação do volume da região de aplicação das tuberculinas, em milímetros.

Para o exame do tipo *ELISA*, Macaé possui menos exames disponíveis, apenas dois: ELISA MPB70 e ELISA MPB83. Para cada ELISA, há o seu valor de densidade ótica e também a indicação das saídas que estão acima do ponto de corte. O ELISA com recombinante MPB70 apresentou seus valores dentro da faixa dinâmica de -0,03 até 0,90 para sua densidade ótica. O recombinante MPB83 apresentou seus valores de densidade ótica em uma faixa dinâmica menor, variando de -0,03 até 0,39.

O conjunto de dados possui os mesmos exames realizados pela UFF (Seção 4.1.1) para os tecidos: Ziehl Neelsen para cultura, a PCR direta sobre a amostra e PCR sobre a cultura da amostra. Tais testes também são considerados como o padrão ouro da fazenda. Os diagnósticos por estes exames foram obtidos da mesma maneira que a fazenda UFF, onde tendo o animal reagido positivamente aos testes cervicais ou ao ELISA, tal animal era sacrificado e os exames de tecido eram feitos para confirmação. Não foram aplicados para esta fazenda exames relativos ao leite.

Tabela 4.5: Relação de exames usados, tipo do dado e seus possíveis valores - Macaé

<b>Campo</b>	<b>Tipo</b>	<b>Valores</b>
<i>PPD Bov</i>	Numérico	[0, 20 ; 19, 40]
<i>PPD Av</i>	Numérico	[-0, 20 ; 8, 40]
ELISA MPB70	Numérico	[-0, 03 ; 0, 90]
ELISA MPB83	Numérico	[-0, 03 ; 0, 39]
Cultura Tecido	Catégorico	Pos; Neg; Inc; '-'
PCR Cultura Tecido	Catégorico	Pos; Neg; '-'
PCR direta Tecido	Catégorico	Pos; Neg; '-'
Anérgicos	Catégorico	Susp; Neg

Por fim, a base da fazenda ainda apresenta informação de quais animais são suspeitos de serem anérgicos. Nesta base, um total de 4 animais são suspeitos e estes animais tiveram a confirmação da doença pelos testes de tecido. A prevalência de tuberculose na fazenda Macaé é de 46% dos animais.

### 4.1.3 Fazenda RS

O conjunto de dados referido por *Fazenda RS* é referente a um conjunto de três propriedades privadas, localizadas nos municípios de Arroio do Meio e Bom Retiro do Sul, estado do Rio Grande do Sul. Apesar dos dados serem provenientes de mais de uma fazenda, estes não estão distribuídos de maneira uniforme por fazenda, tendo uma das fazendas muito mais animais que as demais. Os dados foram obtidos a partir de programas de controle da Tuberculose bovina implementados na região, sendo eles os de registro 02.13.16.002.00.00 pela EMBRAPA, 443235/2014-7 pelo CNPq, e 085/2015 pela Fundect.

A base de dados possui um total de 31 animais, com seus diagnósticos definidos, com a relação da quantidade de animais por diagnóstico exposta na Tabela 4.6. Para esta fazenda, a quantidade de animais com e sem tuberculose bovina está desbalanceada, possuindo a fazenda 26 animais (84% do total) confirmados com tuberculose bovina e 5 animais (16% do total) considerados livres da doença. A base de dados não contém qualquer informação sobre a suspeita ou confirmação

de haver animais anérgicos nas fazendas. Logo, a variável “anergia” em animais tuberculosos não é considerada em RS.

Tabela 4.6: Quantidade de Animais TB+ e TB− para Fazenda RS

Status TB	Ocorrências	Porcentagem
Infectado	26	84%
Não Infectado	5	16%
<b>Total</b>	31	100%

Como nas Fazendas UFF e Macaé, cada animal possui um conjunto de exames realizados, porém, não os mesmos se comparado aos outro conjuntos de dados. A Tabela 4.7 apresenta a lista dos dados disponíveis e a descrição do formato que cada campo da base pode assumir. Para os campos com valores numéricos, a tabela ainda apresenta o intervalo que os contém.

Os valores de PPD Bovina e Aviária são usados para a determinação do TCS e do TCC, representando a variação de volume da região de injeção das tuberculinas, ocorrendo até no máximo 22,30 mm e 16,70 mm, respectivamente.

A fazenda RS possui dois exames de tecido, que são considerados como os testes de padrão ouro. O primeiro teste confirmatório é o exame de cultura das amostras de tecido e o segundo é o chamado *Nested-PCR* [46]. Todos os animais dessa base possuem o exame de cultura já finalizado mas nem todos possuem *Nested-PCR* terminado e portanto, a determinação do diagnóstico foi feita considerando apenas os exames completados de cada animal.

O teste ELISA usado nos animais desta base é o chamado ELISA quimérico, desenvolvido pela EMBRAPA [31]. Como descrito na Seção 2.2.2, o ELISA quimérico usa uma combinação dos antígenos ESAT-6, MPB70 e MPB83. O resultado retornado por este ELISA está no formato de densidade ótica, assim com os demais testes ELISA apresentados. No entanto, o valor numérico mostrado na base de dados de RS não se refere aos valores de densidade ótica, mas a um valor normalizado. A normalização usada para a geração de tais valores é a chamada normalização *NAR*.

### Normalização *NAR*

Diversos fatores, técnicos ou de ambiente, podem influenciar os resultados do exame ELISA. A temperatura ambiente no momento em que o teste é efetuado, a lavagem feita nas placas que recebem o soro, o preparo dos reagentes usados em cada exame e os aparelhos que fazem as leituras dos resultados são fatores que geram uma variabilidade nos resultados do teste ELISA [79]. Diferentes kits de ELISA também podem apresentar diferenças nos resultados gerados [80] [81]. Tal variabilidade é

Tabela 4.7: Relação de exames usados, tipo do dado e seus valores - RS

<b>Campo</b>	<b>Tipo</b>	<b>Valores</b>
<i>PPD Bov</i>	Numérico	[0 ; 22, 30]
<i>PPD Av</i>	Numérico	[0 ; 16, 70]
ELISA	Numérico	[0, 32 ; 21, 50]
Cultura Tecido	Categórico	Pos; Neg; Inc; ‘-’
PCR direta Tecido	Categórico	Pos; Neg; ‘-’

vista em situações práticas, onde uma mesma amostra pode apresentar medidas de densidade ótica diferentes devido a algumas destas variáveis.

A técnica de normalização usada sobre o teste ELISA é a chamada *NAR* (*Normalized Absorbance Ratio*) [79]. A normalização *NAR* é uma técnica que visa diminuir as variações dos resultados do teste ELISA. Pela técnica, devem ser adicionadas nas placas usadas no teste, não somente as amostras a serem analisadas, mas também amostras de grupos de controle e amostras ditas “em branco”. A ideia básica da normalização *NAR* é usar esses componentes para corrigir os valores das amostras testadas.

Os chamados grupos de controle, considerando o problema específico da tuberculose em bovinos, são amostras que sabidamente representam um animal com uma das duas condições de diagnóstico: animal tuberculoso ou animal sadio. Amostras pertencentes ao grupo de controle positivo são amostras referentes a animais que sabidamente estão infectados pelo agente *M. bovis*. Já as amostras pertencentes ao grupo de controle negativo são amostras recolhidas de animais que não sejam tuberculosos. As amostras “em branco” contém apenas o solvente usado na solução em que as amostras de soros dos animais são aplicadas.

O valor final normalizado é calculado em três etapas distintas. A primeira etapa consiste em fazer as leituras de densidade ótica dos animais do grupo de teste, dos animais de grupos de controle e das amostras ditas “em branco”, utilizando instrumentos capazes de detectar os valores de OD também deste último grupo. Sendo assim, a segunda etapa consiste em corrigir as amostras lidas e os grupos de controle subtraindo valores de densidade ótica das leituras “em branco”. A terceira e última etapa consiste em dividir o valor corrigido da amostra pelo valor corrigido do grupo de controle. Caso um animal ou grupo de controle testado possua mais de uma amostra testada, a normalização é feita considerando-se o valor médio deste grupo. A fórmula da *NAR* é dada pela equação 4.1.

$$NAR = \frac{(\text{Média Amostra} - \text{Média Branco})}{(\text{Média Controle} - \text{Média Branco})} \quad (4.1)$$

Uma amostra testada no ELISA e normalizada por *NAR* tem seu valor final definido pelo grupo de controle. Logo, a normalização pode ser feita tanto sobre o grupo de controle positivo, como pelo grupo negativo. A fazenda RS possui em seus dados, os valores de densidade ótica para os dois grupos controle.

A faixa dinâmica do resultado do teste ELISA, apresentada na Tabela 4.7, é dada considerando os valores normalizados gerados pela técnica de *NAR* tanto para o grupo de controle positivo quanto para o grupo de controle negativo, presentes no conjunto de dados.

#### 4.1.4 Preparação dos dados

Descritas as bases de dados, estas ainda precisam ser tratadas, antes de terem seus valores passados aos modelos neurais e serem usados nas análises. A fase de preparação visa encontrar e tratar inconsistências presentes nos dados e em seguida formatá-los, para o seu uso nos modelos neurais.

##### Busca por inconsistências

O primeiro passo consistiu em buscar possíveis inconsistências nas bases de dados, sendo o primeiro o de bovinos com valor negativo para a variação do volume da região em que foram aplicadas as PPDs Bovina ou Aviária. Tais casos indicam que a região de aplicação da tuberculina diminuiu seu tamanho, ao invés de apresentar um inchaço e um conseqüente aumento de volume. O PNCEBT [3], considerando os testes cervicais, permite que apenas a diferença entre as regiões de aplicação das duas tuberculinas seja negativa, como visto na Tabela 2.2 para o TCC.

O segundo caso de inconsistência se refere a amostras com dados faltantes para os testes cervicais, impossibilitando a determinação do diagnóstico para o TCS e TCC.

Os registros de animais nas bases de dados, com alguma das inconsistências encontradas, foram removidos e desconsiderados para o uso pelos modelos neurais. A Tabela 4.8 indica a relação dos animais que foram removidos e dos que permaneceram em cada fazenda, separando-os em: Animais anérgicos, animais tuberculosos e animais sadios.

Tabela 4.8: Relação de animais mantidos e animais removidos por fazenda

<b>Animal</b>	<b>Situação</b>	<b>UFF</b>	<b>Macaé</b>	<b>RS</b>
Anérgico	Utilizados	3 (4%)	4 (8%)	0 (0%)
	Removidos	4 (5%)	0 (0%)	0 (0%)
TB +	Utilizados	20 (26%)	23 (46%)	25 (81%)
	Removidos	4 (5%)	0 (0%)	1 (3%)
TB –	Utilizados	23 (30%)	22 (44%)	5 (16%)
	Removidos	23 (30%)	1 (2%)	0 (0%)
<b>Total</b>	<b>Utilizados</b>	<b>46 (60%)</b>	<b>49 (98%)</b>	<b>30 (97%)</b>
	<b>Removidos</b>	<b>31 (40%)</b>	<b>1 (2%)</b>	<b>1 (3%)</b>

Para a fazenda UFF, encontrou-se apenas casos em que os valores de injeção de tuberculina eram valores negativos, tanto para a tuberculina bovina quanto a tuberculina aviária. Para essa fazenda, não foram encontrados dados não preenchidos. Foram removidos de UFF 31 animais (40%) de um total de 77 registrados na fazenda, dos quais, 4 eram anérgicos, 4 eram positivos para tuberculose apenas e os demais 23 (30% do total) eram animais livres da doença. A fazenda Macaé, que possui um total de 50 animais, apresentou apenas 1 animal (2%) removido, por este apresentar valor negativo para as tuberculinas injetadas. O bovino descartado da base não era um anérgico. Já na fazenda RS, não foram encontrados valores inconsistentes para os campos de PPD Bovina e PPD Aviária. No entanto, foi encontrado 1 registro de bovino (3% do total) sem valores referentes aos seus testes cervicais, sendo impossível a determinação do diagnóstico por estes exames. O animal removido não era anérgico.

### Formatação dos dados

Selecionadas as amostras que seriam efetivamente usadas nos experimentos, o próximo passo consistiu em definir um formato para o seu uso nos modelos neurais a serem desenvolvidos. Decidiu-se por usar variáveis de entrada com valores categóricos, em formato similar ao apresentado na Tabela 4.1. Foram considerados os resultados de cada um dos exames e seus possíveis valores como apenas: Positivo, Negativo ou Inconclusivo. Tais valores foram convertidos em uma representação numérica, para as redes neurais. A Tabela 4.9 mostra a relação dos valores e seus significados.

Tabela 4.9: Representação do diagnóstico na Rede Neural

<b>Significado</b>	<b>Representação na Rede</b>
Positivo	1
Inconclusivo	0
Negativo	-1

O segundo ponto foi a escolha dos exames a serem aplicados aos modelos neurais, que foi baseado na existência dos resultados ou na possibilidade de calculá-los a partir dos dados disponíveis. Os exames selecionados seguem abaixo:

- TCS: Resultado do TCS
- TCC: Resultado do TCC
- ELISA: Resultado dos teste ELISA
- Leite: Resultado dos testes de amostras de leite
- Status: Diagnóstico para Tuberculose Bovina
- Anérgico: Diagnóstico para Anergia

Para as entradas TCS e TCC, todas as fazendas não apresentavam seus resultados finais, mas sim seus valores de PPD Bovina e PPD Aviária e estes foram usados para gerar os diagnósticos, seguindo a norma brasileira. Em seguida, os resultados foram passados ao formato definido anteriormente na Tabela 4.9.

O teste ELISA, para todas as fazendas, também possui valores capazes de serem transformados em um diagnóstico. Para as fazendas UFF e Macaé, cada teste ELISA possui o seu respectivo valor de densidade ótica com a indicação de quais valores estavam acima do ponto de corte. O resultado final foi gerado através de um *OU-Lógico*, bastando que apenas um dos testes retornasse positivo, caso contrário o resultado era considerado negativo.

Já para a fazenda RS, a definição de resultado do ELISA foi feita de modo diferente, uma vez que seus valores não são os valores brutos de densidade ótica, mas sim valores normalizados por *NAR*. O primeiro passo consistiu em decidir qual grupo de controle usar: grupo de controle Positivo ou Negativo. Uma abordagem de decisão seria o estudo feito por LILENBAUM [82], que adotou o critério de usar animais livres de tuberculose como grupo de controle e considerar que valores de ELISA acima dos valores deste grupo são considerados doentes. Esta abordagem busca identificar o maior número possível de bovinos positivos, aumentando a taxa

de Sensibilidade mas também aceitando um aumento na taxa de False Alarme pelo ELISA. Um outro caso, escolhendo-se o grupo de controle de animais positivos, há um aumento da Especificidade do ELISA e um aumento na taxa Falso Negativo. A não identificação de bovinos doentes em um rebanho pode significar a disseminação da doença no mesmo e por este motivo, foi escolhido o grupo de controle negativo para o exame ELISA de RS.

Com o grupo de controle escolhido, aplica-se os valores de OD desse grupo na equação 4.1 e cada amostra é normalizada a partir destes valores. A Tabela 4.10 trás a interpretação para os possíveis resultados normalizados. No caso, um valor normalizado superior a 1 ( $NAR > 1$ ) indica que o animal possui um valor para ELISA maior que os animais negativos para tuberculose, indicando um diagnóstico positivo da doença. Para um valor igual ou inferior a 1 ( $NAR \leq 1$ ), a interpretação deste resultado é que seu valor de densidade ótica é, no máximo, igual ao valor de um animal sem tuberculose bovina, indicando que o animal não está infectado.

Tabela 4.10: Interpretação dos valores normalizados de ELISA

Valor $NAR$	Situação	Diagnóstico
$\leq 1$	OD Amostra $\leq$ OD Controle	Negativo
$> 1$	OD Amostra $>$ OD Controle	Positivo

Para o teste de leite, como apenas a fazenda UFF possui tal exame, esta variável foi desconsiderada nas fazendas Macaé e RS. Nas amostras de UFF que não possuem o leite feito, o resultado foi considerado como inconclusivo. Para os bovinos em que o teste foram realizados, o critério usado foi o de um *OU-Lógico* para os testes de Ziehl Neelsen, PCR sobre a cultura do leite e a PCR sobre amostra direta do leite para determinar seu diagnóstico, seguindo o mesmo critério adotado para no ELISA de UFF e Macaé.

O campo Status define se o animal possui ou não Tuberculose Bovina e este campo usa os exames de tecido para a determinação de seu valor, que pode ser Positivo ou Negativo. Novamente, o critério usado foi um *OU-Lógico* aplicado sobre todos os testes de tecido das fazendas. Os exames de tecido são do tipo *post-mortem* e portanto, dependem que o animal seja sacrificado para que o teste seja feito, sendo estes realizados apenas em casos de extrema de necessidade e somente em animais que apresentaram um diagnóstico positivo para qualquer um dos exames *ante-mortem* (testes cervicais, ELISA e leite). O campo Status é definido como positivo ou negativo, dependendo do diagnóstico obtido pelos exames de tecido (presença ou não de *M. bovis*) e os animais que apresentaram nenhuma reação aos exames *ante-mortem*, são considerados negativos.

O campo *Anérgico* pode assumir apenas os valores de positivo e negativo. Para o caso do animal ser considerado suspeito de ser anérgico, o valor deste campo é definido como positivo, caso contrário, negativo. Após os passos descritos nesta seção, os dados foram considerados prontos para uso e análise.

## 4.2 Adição de exames aos testes cervicais

Desde a criação do PNCEBT, em 2001, o programa tem gerado resultados positivos e conseguido reduzir a prevalência da tuberculose bovina. Estudos conduzidos no período de 1989 ao ano de 1998, antes da criação do programa, indicavam uma prevalência de aproximadamente 1,3% do rebanho nacional infectado [26], porém há regiões do Brasil que têm atingido valores próximos da erradicação da doença. O estado do Mato Grosso estima que, em 2009, a prevalência de tuberculose bovina em animais e rebanhos abatidos no estado era de aproximadamente 0,007% do total de animais do estado [52].

Apesar dos avanços, os exames hoje definidos pelo PNCEBT não apresentam resultados que garantam a total identificação da doença em um rebanho. Estudo conduzido por MARASSI [15] indica que dentre os exames possíveis de serem usados no diagnóstico da tuberculose, nenhum deles consegue identificar todos os animais, em todos os estágios da infecção. Contudo, o mesmo estudo indica que o uso destes exames em conjunto aumenta a eficiência na identificação dos animais infectados.

Este experimento, assim como feito em [15], testa a abordagem multidisciplinar aplicada no diagnóstico de tuberculose bovina. O objetivo do experimento é avaliar o impacto que a adição de novos exames, aos testes já preconizados pelo PNCEBT, gera na identificação da doença, sendo medido pelas medidas de: Sensibilidade, Especificidade, Acuidade e valor SP.

Técnicas de Aprendizado de Máquina (*Machine Learning*), como por exemplo as Redes Neurais Artificiais, são aplicadas como apoio ao diagnóstico em diferentes problemas da área médica [8], sendo assim possível utilizar tais técnicas no cenário da Tuberculose Bovina.

Uma abordagem neural é proposta para avaliar o acréscimo de exames e o seu uso em conjunto na identificação da tuberculose bovina. Para este experimento, foram desenvolvidas Redes Neurais Artificiais, usadas como classificadores.

Nesta seção são descritos os modelos neurais desenvolvidos, a topologia escolhida e o treinamento aplicado aos modelos neurais.

### 4.2.1 Os modelos neurais

No Brasil, a erradicação da tuberculose bovina ainda não foi alcançada devido a grande dificuldade em conseguir identificar todos os animais tuberculosos presentes nos rebanhos do país. Por esta característica, o problema pode ser considerado como um problema de classificação. Especificamente para tuberculose bovina, a partir de um conjunto de informações sobre o animal, busca-se determinar se o mesmo está doente. Pela descrição do problema, os modelos neurais para este experimento foram idealizados como classificadores de tuberculose bovina. Estes classificadores não foram utilizados com o intuito de se avaliar sua capacidade de classificação de animais com e sem tuberculose, mas para avaliar o impacto de novos exames adicionados ao já preconizados pela legislação brasileira.

Os dados disponíveis, descritos na Seção 4.1, após sua formatação, são utilizados como as informações de entrada dos modelos neurais. Cada modelo recebe um conjunto de testes para tuberculose e seus resultados, colocados no formato descrito na Tabela 4.9. As saídas dos modelos neurais também são colocados no mesmo formato dos dados de entrada, com exceção que a saída não apresenta o resultado inconclusivo.

Cada uma das fazendas teve um conjunto de modelos neurais desenvolvidos separadamente das demais, devido a alguns fatores que as distinguem. O primeiro deles diz respeito ao ambiente das fazendas, uma vez que UFF pertence a uma universidade e é utilizada para fins de pesquisa e ensino, não possuindo interesses econômicos como as fazendas de gado leiteiro e gado de corte. Logo, quando o animal é diagnosticado como doente, ele é sacrificado. Esse fato faz com que essa fazenda seja considerado um cenário pouco provável.

Já as fazendas Macaé e RS são privadas, utilizadas com fins de lucro, e é encontrada uma resistência por parte dos proprietários a sacrificar seus animais, sendo sacrificados somente na confirmação total da doença. Essas fazendas são um cenário real.

Outro fator que difere a fazenda UFF das demais está no controle e rigor da execução dos exames, principalmente para os exames cervicais. A fazenda UFF, por possuir um cunho de pesquisa e contar com um corpo técnico de veterinários, gera seus resultados com boa precisão. Nas demais fazendas, a medição nem sempre é feita por profissionais responsáveis e autorizados, o que pode acarretar em um aumento do erro sistemático das amostras colhidas. O último fator está no ponto das três fazendas possuírem exames diferentes entre si, com somente a UFF contendo o exame do leite e todas as fazendas possuindo um conjunto diferente de exames ELISA.

A Tabela 4.11 trás a relação dos modelos neurais desenvolvidos para cada fa-

zenda. As fazendas UFF e Macaé, por possuírem animais anérgicos, serão treinadas em dois casos distintos: com e sem a presença de animais anérgicos. Outra distinção dos modelos diz respeito à fazenda UFF, onde manteve-se o exame do leite na análise, sendo este considerado apenas no cenário de sua fazenda. Para efeitos de comparação entre as fazendas, tal exame não será considerado. Há portanto, um total de 14 modelos neurais usados para este experimento.

Tabela 4.11: Relação dos modelos neurais por fazenda

Fazenda	Anérgicos	Exames por rede
UFF	COM SEM	TCS e TCC
		TCS, TCC e ELISA
		TCS, TCC e Leite
		TCS, TCC, ELISA e Leite
Macaé	COM SEM	TCS e TCC
		TCS, TCC e ELISA
RS	SEM	TCS e TCC
		TCS, TCC e ELISA

## 4.2.2 Topologia dos modelos

A topologia de uma rede neural se refere ao modo como seus neurônios são conectados na mesma, e este é um fator muito importante para o funcionamento e aprendizado da rede. Para a definição das topologias dos modelos neurais desenvolvidos neste experimento, três diferentes parâmetros foram considerados: quantidade de camadas da rede, suas funções de ativação e a quantidade de neurônios utilizados.

O primeiro parâmetro averiguado é a quantidade de camadas que a rede neural possui. Para este estudo decidiu-se usar uma rede Perceptron Multicamadas (*Multi-Layer Perceptron*) com apenas uma camada escondida de neurônios. Logo, os modelos neurais possuem três camadas distintas: uma camada de entrada, uma camada intermediária (ou camada escondida) e uma camada de saída. A escolha de MLP com apenas uma camada escondida foi feita baseada no fato de, caso a rede neural possuísse apenas as camadas de entrada e saída, o modelo neural seria um classificador linear e conseqüentemente, não seria capaz de resolver problemas de maior complexidade e não-linearmente separáveis [62]. A maioria dos problemas podem ser resolvidos por modelos neurais com apenas uma camada interna de neurônios e o uso de mais camadas internas está ligada a fatores como: a altíssima complexidade do problema e grande quantidade de amostras disponíveis para o aprendizado do modelo [67].

O segundo parâmetro analisado para a topologia são as funções de ativação que cada neurônio deverá utilizar. Escolheu-se para os neurônios, tanto da camada escondida quanto da camada de saída, a função Tangente Hiperbólica, que é uma das mais comumente usadas por técnicas de Aprendizado de Máquina [83]. O primeiro fator para seu uso é o fato desta ser uma função não linear e podendo ser aplicada, considerando casos de classificação, em problemas linearmente e não linearmente separáveis. O segundo fato é dado pelo teorema de aproximação universal [84], que dita que toda função contínua, onde são mapeados intervalos de números reais de entrada a algum intervalo de números reais de saída, pode ser arbitrariamente aproximada com precisão por uma rede perceptron multicamadas com somente uma camada interna. Esse resultado é válido para as funções Sigmoides [63], e também para a função Tangente Hiperbólica, uma vez que esta é uma variante da primeira. O problema da tuberculose bovina, como proposto, por ser dado como um problema de classificação binária, com seus diagnósticos formatados para 1 e -1 (Tabela 4.9), tem-se que Tangente Hiperbólica é aplicável ao problema. A Figura 4.1 apresenta o formato aproximado definido para os modelos neurais, com exceção do número de neurônios na camada escondida, com o exemplo de um modelo neural que recebe como entrada os exames TCS, TCC, ELISA e o teste de leite.

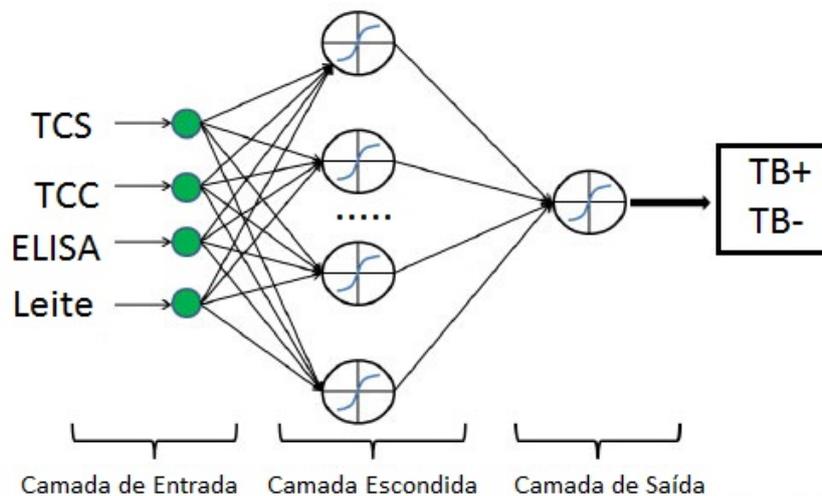


Figura 4.1: Exemplo de Topologia de modelo neural desenvolvido

O terceiro parâmetro diz respeito à quantidade de neurônios da camada escondida das redes neurais, o qual é um fator crucial para o seu desempenho. Busca-se evitar que a escolha do número de neurônios resulte em situações de *Underfitting* e *Overfitting*. Cada modelo neural possui um número ótimo de neurônios, que otimiza seu desempenho, considerando medidas de performance de Sensibilidade, Especificidade e valor SP. A determinação deste valor foi feita treinando-se todos os modelos, como descrito na Seção 4.2.3, e variando o número de neurônios para a camada interna no intervalo de 1 a 10. O valor SP dos grupos de teste dos modelos

foi a métrica de performance usada para a escolha da quantidade de neurônios, sendo escolhida a quantidade que apresentasse o maior valor para a medida.

### 4.2.3 Treinamento

No contexto específico das redes neurais, a fase de treinamento é fase onde há a determinação e o ajuste dos pesos das sinapses. Dado que as amostras nas bases de dados possuem um valor alvo disponível, supõe-se o uso de treinamento supervisionado para os modelos neurais do experimento. O resultado alvo de cada amostra foi dado pelo campo “Status” (Tabela 4.9), após a preparação dos dados brutos.

Tendo escolhido se usar uma *MultiLayer Perceptron* e o paradigma de aprendizado supervisionado, o algoritmo utilizado para o treinamento das redes neurais foi o *Resilient Backpropagation*. Um dos fatores para sua escolha se deu por este ser um dos algoritmos de atualização de pesos mais rápido e por ser adaptativo [69], requerendo menos parâmetros para serem configurados. Escolhido o algoritmo, alguns dos seus parâmetros ainda precisam ser ajustados.

O primeiro parâmetro definido para o treinamento foi o critério de parada, onde esse somente seria encerrado quando o número máximo de épocas escolhida para o treinamento fosse atingido. Não foi usado nenhum critério de Parada Antecipada (*Early Stopping*) [85], sendo esta uma técnica usada para se evitar o *Overfitting* no treinamento de um modelo. A Parada Antecipada é definida, por exemplo, para que o treinamento do modelo seja interrompido quando o mesmo apresente um aumento no erro para o grupo de validação, indicando um possível estado de *Overfitting*.

No entanto, a técnica de *Early Stopping* possui limitações, podendo fazer com que o algoritmo pare seu treinamento em uma região de mínimo local para seu grupo de validação. A Figura 4.2 apresenta um exemplo hipotético desta situação, onde seu eixo horizontal relaciona as épocas de treinamento do modelo neural e seu eixo vertical, o erro de validação do modelo. A figura apresenta a evolução do erro de validação pelas épocas de treinamento e é visto que este exemplo possui diferentes valores de mínimo local (destacados em vermelho) e um mínimo global (destacado em verde). Pela Figura 4.2, percebe-se que o valor de mínimo global é encontrado apenas em uma fase avançada do treinamento do modelo, próximo da metade do total de épocas transcorridas, e que o critério de *early stopping* sendo definido para uma parada precoce, levará o modelo a não terminar seu treinamento na região de mínimo global.

Porém, mesmo sendo desconsiderada a parada antecipada e sendo o modelo treinado no total de épocas definidas, usa-se a estratégia chamada *Save the best* para escolher a melhor configuração de valores para os pesos sinápticos do modelo neural, dentre todas as configurações de valores apresentadas em cada época de

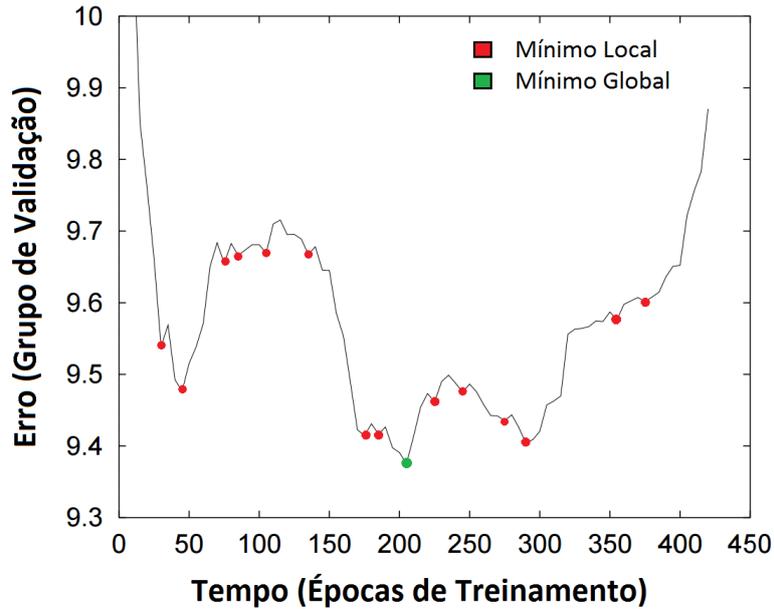


Figura 4.2: Exemplo de evolução do erro de validação de modelo neural

treinamento. A medida de performance considerada por todo o processo foi o valor SP.

O segundo parâmetro ajustado é o de épocas de treinamento, que é analisado em conjunto com o seu critério de parada. Uma época na etapa de treinamento corresponde a uma iteração de atualização dos pesos sinápticos da rede neural e da propagação dos dados de validação. Ao final de cada época, o modelo neural é avaliado de acordo com a medida escolhida (Ex: Erro médio quadrático e valor SP) para estes dados. Para o algoritmo, este parâmetro determina a quantidade máxima de épocas que o treinamento poderá alcançar, que foi fixado em 10000 épocas. A escolha por esse valor foi feita para aumentar a chance de convergência para uma região de mínimo global, no entanto, um valor alto para épocas de treinamento tem como consequência o aumento do tempo de treinamento dos modelos.

Outro critério considerado foi o chamado *Batch Size* ou tamanho do lote, que é um conceito relacionado com o modo de treinamento dos modelos neurais, havendo dois modos básicos utilizados nesta etapa: Modo Sequencial e Modo Batelada (*Batch Mode*) [6]. No modo sequencial, a cada época do treinamento, apenas uma amostra é apresentada e os pesos do modelo neural são ajustados. Já no modo Batelada, para cada época de treinamento é apresentado um conjunto de amostras dos dados e não mais uma amostra unitária. O *Batch Size* é a variável que determina quantas amostras são passadas para cada época de treinamento e pode-se considerar que o modo Sequencial de treinamento é um caso particular do modo Batelada, onde o tamanho do lote é igual a uma amostra apenas.

Para o experimento foi escolhido o modo de treinamento em Batelada. O ta-

manho de *Batch Size* foi definido de acordo com os dados que eram utilizados e foi usada a mesma quantidade de animais TB+ e TB− em cada época de treinamento. Tal escolha foi feita para que o treinamento fosse feito de forma balanceada, e a rede neural aprendesse de maneira igual os dois tipos bovinos, TB+ e TB−. Para a determinação do tamanho do lote, verifica-se entre as duas classes de animais (TB+ e TB−), qual delas possui a menor quantidade de amostras e então a classe com mais amostras terá, para cada época de treinamento, a mesma quantidade de animais da outra classe. Em cada época, a menor classe tem todas as suas amostras passadas para o treinamento, enquanto a segunda classe de animais tem suas amostras definidas randomicamente.

Para a divisão dos grupos de treinamento, validação e teste, considerou-se primeiramente que os grupos de validação e teste seriam os mesmos devido à escassez de dados nas bases disponíveis. A quantidade de dados atribuída a cada um dos grupos foi definida de acordo com a fazenda usada. Para as fazendas UFF e Macaé, foi definido que seus grupos de treinamento e teste/validação receberiam 50% do conjunto de dados cada um, uma vez que estas fazendas possuem maior quantidade de dados e serem mais balanceadas que a fazenda RS. Já a fazenda RS possui aproximadamente 15 amostras a menos que as demais fazendas, o que para o problema é uma grande diferença. Se em RS fosse adotada a mesma proporção usada nas demais fazendas, isso poderia comprometer o aprendizado dos modelos e por isso, a proporção adotada foi de 75% dos dados para treinamento e 25% dos dados para teste/validação. A Tabela 4.12 mostra a relação de dados pelos grupos de treinamento e teste/validação.

Tabela 4.12: Proporção dos dados para cada grupo

<b>Fazenda</b>	<b>Treinamento</b>	<b>Teste/Validação</b>
UFF	50%	50%
Macaé	50%	50%
RS	75%	25%

Juntamente com o treinamento, foi definido o uso da técnica de validação cruzada. Para as três fazendas, os dados foram divididos usando o método *k-Fold*, em um total de 12 blocos de dados, onde o tamanho dos blocos era definido pelo total de dados de cada fazenda. Esses blocos foram distribuídos aleatoriamente de modo a atender as proporções definidas para o conjunto de treino e teste de cada fazenda (Tabela 4.12), ou seja, para as fazendas UFF e Macaé 6 blocos eram usados para treinamento e os demais para validação e teste. Em RS foi usado um total de 9 blocos de dados para treinamento e apenas 3 blocos para validação e teste.

O método *k-fold* de Validação Cruzada, como visto na Seção 3.2.4, determina que os blocos de dados sejam sorteados aleatoriamente para os respectivos grupos de treinamento, validação e teste. Para o experimento, foi definido que seriam realizados 50 sorteios destes blocos, respeitando a proporção estabelecida. Na prática, são treinadas 50 redes neurais diferentes, com diferentes conjuntos de treinamento e validação/teste. Escolheu-se fazer 50 sorteios para aumentar as combinações de blocos de dados usados. Para diminuir a chance que estas 50 redes não convergissem para um mínimo local, 100 inicializações de pesos foram feitas para cada uma delas. Dentre todas as inicializações feitas para a mesma combinação de blocos de dados, a que apresentassem o maior valor SP de teste, era selecionada.

### 4.3 Repetibilidade

No método científico existem dois conceitos muito importantes para a medição e a confiabilidade dos resultados de um experimento: a Repetibilidade e a Reprodutibilidade. A Repetibilidade busca a confirmação de um resultado obtido através da repetição do experimento, sob as mesmas condições definidas em todas as suas repetições. As condições a serem mantidas podem se referir ao método de medição utilizado, aos equipamentos utilizados, ao observador ou ainda ao local utilizado, no entanto, não é necessário que os dados utilizados na Repetibilidade sejam os mesmos. Já a Reprodutibilidade busca aproximar o resultado do experimento original com o resultado obtido para um novo experimento, similar ao primeiro, porém com variações em suas condições, tais como mudanças no método de medição, nos instrumentos de medida e no operador [86].

A presente seção apresenta um segundo experimento, complementar ao primeiro, explicitado na Seção 4.2, e seu objetivo é novamente verificar o impacto na identificação da tuberculose com a adição de novos exames aos testes cervicais e validar os resultados encontrados no primeiro experimento. A proposta é de usar a Repetibilidade no primeiro experimento. Para isso, são mantidas as mesmas condições do experimento anterior, como a mesma topologia das redes, o mesmo algoritmo de treinamento, a proporção dos dados para os grupos de treinamento, validação e teste, porém sendo utilizados outros dados.

No primeiro experimento, as amostras que compunham o grupo de teste dos modelos eram as mesmas definidas para o grupo de validação das redes neurais, porém, a proposta deste experimento é utilizar os dados de fazendas em modelos neurais desenvolvidos com dados de outra fazenda.

Nesta seção são descritos como os dados são distribuídos para o experimento, os modelos neurais escolhidos e como estes são usados.

### 4.3.1 Uso dos dados e escolha dos modelos neurais

Para o experimento foi decidido inicialmente como usar os dados disponíveis na Repetibilidade. Primeiramente foi considerado o cenário de cada fazenda no problema da Tuberculose Bovina. A fazenda UFF pode ser considerada como um cenário improvável de ser encontrado em uma aplicação prática. A fazenda possui um ambiente diferente das fazendas que efetivamente são alvo do programa nacional de controle da doença, como por exemplo as fazendas Macaé e RS.

Foi considerado ainda o fator regional e geográfico dos bovinos. Tal fator não atrapalha a identificação da doença nos animais mas pode afetar, por exemplo, o rastreamento da origem da infecção do animal. Como mostrado em [87], o conteúdo genético de agentes *M. bovis* encontrados em bovinos de diferentes regiões do estado do Rio Grande do Sul (RS) eram diferentes entre si e ainda verificou-se que em [88] que o material genético de *M. bovis* de animais da mesma fazenda tendem a ser mais similares do que de outras regiões.

Para o uso dos dados, foi decidido que fazenda RS seria usada como grupo de teste nos modelos neurais desenvolvidos para fazendas de UFF e Macaé. Nesta escolha foi considerada a geografia das três fazendas, onde UFF e Macaé são mais próximas entre si e a fazenda RS é mais afastada das demais. Também foram considerados, nas fazendas, os cenários que as caracterizam, onde UFF é um cenário pouco provável e Macaé um cenário comum aos casos de tuberculose.

Determinadas as funções de cada fazenda, o passo seguinte foi a escolha dos modelos neurais testados. A escolha feita foi de usar apenas os modelos que contivessem exames pertencentes também à fazenda RS. Não foram utilizados os modelos neurais da fazenda UFF que consideravam o exame de leite, e os modelos contendo os exames TCS, TCC e ELISA foram mantidos, sendo estes com e sem a presença de anérgicos no treinamento. A Tabela 4.13 exhibe a relação dos modelos neurais em que foram propagados os dados de RS.

Tabela 4.13: Relação dos modelos neurais usados para a propagação de RS

Valor Fazenda	Anérgicos	Modelos Neurais
UFF	SEM	TCS e TCC
	COM	TCS, TCC e ELISA
Macaé	SEM	TCS e TCC
	COM	TCS, TCC e ELISA

### 4.3.2 Redes de Operação

Neste estudo, a chamada “Propagação” é feita a partir de dados de entrada, recebidos pela rede neural, que são transmitidos por suas sinapses e neurônios artificiais até a camada de saída, gerando uma resposta à entrada recebida. Esse processo é feito, por exemplo, nas etapas de validação e teste do modelo neural. Para o cenário descrito em nossos experimentos, a propagação é dada pela apresentação dos exames de uma certa amostra para os modelos neurais e estes retornam o diagnóstico para TB.

Para cada modelo neural considerado, foi usada a técnica de validação cruzada e assim, foram geradas um total de 50 redes neurais diferentes para cada modelo, onde cada uma foi treinada com uma configuração diferente dos dados de treinamento, validação e teste (Seção 4.2). No entanto para este experimento de Repetibilidade, foi decidido propagar a fazenda RS não para todas as redes neurais geradas na validação cruzada, mas apenas para as chamadas Redes de Operação.

O conceito de Redes de Operação, como visto na Seção 3.1, é um conceito usado no contexto de Redes Neurais Artificiais. Em casos gerais, a escolha de uma rede de operação deve levar em consideração as diferentes amostras que ela pode receber de um cenário real, quando colocada no ar para atuação. Redes Neurais capazes de retornar o mais corretamente possível suas saídas, para a maior parte dos tipos de entradas encontradas, são boas candidatas para serem selecionadas para operação. Considerando especificamente a Tuberculose Bovina, existe uma variedade de entradas para os modelos neurais e casos relacionados à doença, como por exemplo animais com exames de testes cervicais negativos mas positivos para tuberculose ou situações de falso positivo para certos exames. Uma rede de operação aplicada à Tuberculose busca identificar corretamente a maior quantidade e variedade de casos.

As fazendas usadas nos experimentos apresentam limitações para a escolha das redes de operação. O principal fator é a pequena quantidade de amostras que cada uma das fazendas possui, como visto na Tabela 4.8. Uma rede de operação deveria idealmente ser avaliada apenas com dados antes nunca vistos por ela, como em um cenário prático onde esta receberá dados desconhecidos, por exemplo, originadas de diferentes fazendas. No entanto, para os dados disponíveis, os grupos de teste definidos para os modelos neurais (Tabela 4.12) são grupos pouco representativos ao problema da Tuberculose e não contém todas as variações ou casos possíveis da fazenda. Como por exemplo, os animais anérgicos presentes em UFF e Macaé, presentes em pequenas quantidades, os grupos de teste dos modelos neurais conteriam apenas metade da quantidade desses animais. Para contornar este fato, a escolha das Redes de Operação foi feita considerando-se o conjunto de treinamento e teste juntos, ou seja, para cada rede neural gerada na validação cruzada foram propagados

todos os animais da fazenda.

Para cada modelo neural de cada fazenda, todas as redes neurais geradas pela validação cruzada tiveram seus respectivos conjuntos de treinamento e teste totalmente propagados. O valor SP foi a medida de performance definida para escolher as redes de operação, sendo a rede neural com maior valor a escolhida.

Com as redes neurais de operação selecionadas, as amostras da fazenda RS são propagadas nestas redes. Neste experimento são consideradas como medidas de avaliação e performance das redes neurais as medidas de Sensibilidade, Especificidade, Acuidade e valor SP.

## 4.4 Otimização dos testes cervicais

O programa nacional de controle a tuberculose define como testes oficiais para a identificação da doença, os testes cervicais: TCS e TCC. Além dos testes, o PNCEBT determina os valores que definem os diagnósticos para os mesmos, como visto na Seção 2.2.1.

Os exames porém, não tem mostrado capacidade de cobrir a total identificação da tuberculose em rebanhos. DOMENECH *et al.* [16] expôs a limitação dos testes cervicais aplicados em rebanhos de diferentes países. O TCS, por exemplo, identificou corretamente 91% dos animais TB+ de um rebanho na Espanha e 86% na Irlanda do Norte. O mesmo estudo ainda mostrou que o TCC apresentou uma sensibilidade de 91% em rebanhos da Grã Bretanha, 55% na Irlanda do Norte e 93% para a República da Irlanda [16]. No Brasil, os testes intradérmicos também apresentam limitações. LILENBAUM *et al.* [82], analisou animais de 13 rebanhos diferentes, e os testes intradérmicos utilizados apresentaram uma Sensibilidade de aproximadamente 87%, com todos animais analisados juntos.

A Anergia em bovinos é um fator que influencia a taxa de falsos negativos dos testes intradérmicos, uma vez que tal característica é dada pela não reação do animal, mesmo que este esteja doente (Seção 2.2.6). Porém, a anergia em rebanhos é pouco encontrada e isso representa um baixo número de casos deste tipo de animal [50], indicando-se então que o problema da identificação da Tuberculose pela norma brasileira não é causada apenas pelos anérgicos.

A proposta deste experimento é avaliar a possibilidade de reajuste nas normas definidas pelo PNCEBT para os testes cervicais, desejando-se aumentar a Sensibilidade, e minimizar a perda na Especificidade dos exames.

Os ajustes são aplicados separadamente para cada exame, TCS e TCC. Por fim, a norma modificada e otimizada é comparada com a norma definida pelo PNCEBT, e são avaliados os impactos de tais mudanças.

#### 4.4.1 Ajuste da norma oficial

A norma definida pelo PNCEBT é caracterizada pela definição de três possíveis resultados para os testes cervicais: Negativo, Inconclusivo ou Positivo. No entanto, ao se considerar o real estado de infecção de um animal por tuberculose bovina, apenas dois casos são possíveis: Animal infectado ou animal livre do agente *M. bovis*. Sendo assim, o primeiro ajuste considerado é definir como possíveis resultados dos testes cervicais apenas, positivo ou negativo para tuberculose. Com esta nova definição, o resultado dito como inconclusivo é descartado por esta nova norma.

Os diagnósticos dados pelo PNCEBT são gerados através de pontos de corte definidos para o TCS (Tabela 2.1) e o para o TCC (Tabela 2.2). Para ambos, estão definidos dois pontos de corte que delimitam os três resultados possíveis. No ajuste proposto, uma vez que há apenas dois resultados para os exames, é considerado apenas um único ponto de corte que separa os dois diagnósticos.

A norma proposta para os testes intradérmicos usam os mesmos dados usados pela normal oficial, no caso, os valores para as injeções de tuberculina bovina e aviária. Neste experimento, são novamente selecionados apenas os bovinos sem valores negativos para estas tuberculinas. Seus testes de tecido são considerados os testes de padrão-ouro e definitivos para o diagnóstico da doença, tal como descritos na Seção 4.1.4.

#### Determinação do novo corte

Os pontos de corte definidos pela norma oficial são valores fixos e que são usados sem alteração em qualquer rebanho em que se queira verificar a presença de tuberculose. Na proposta deste experimento, o novo ponto de corte é calculado de acordo com a amostra de dados analisadas, não sendo mais um valor fixo por exame.

O novo ponto de corte busca dividir o conjunto de bovinos em animais positivos e animais negativos para tuberculose, mantendo um equilíbrio entre os valores de Sensibilidade e Especificidade. O valor SP é uma medida que possui esta propriedade de equilíbrio (Seção 3.2.3) e por isso, decidiu-se basear a escolha de um ponto de corte no valor SP. O corte associado ao maior valor SP é o ponto considerado ótimo neste experimento.

Para o TCS, o ponto de corte é calculado a partir dos valores de tuberculina bovina ( $\Delta_{Bov}$ ) de uma amostra de dados. O TCC é calculado de maneira análoga ao TCS, porém é usada a diferença entre a tuberculina bovina e a aviária ( $\Delta_{Bov} - \Delta_{Av}$ ). Considere que a variável que  $\Delta$  seja os valores das tuberculinas, dependendo do exame que esteja sendo considerado (TCS ou TCC).

A ideia básica do processo é variar o ponto dentro de intervalo definido e, o valor

com maior SP é escolhido como o ponto de corte ótimo do teste cervical analisado. O ponto de corte baseado em SP é determinado em 4 etapas distintas.

O primeiro passo consiste em verificar qual a faixa dinâmica ou intervalo em que estão contidos os valores de  $\Delta$  da amostra analisada. Esta faixa é definida entre o menor e o maior valor de  $\Delta$  da amostra. O segundo passo consiste em definir os diferentes pontos de corte que serão testados. Para este passo, foi decidido dividir a faixa dinâmica em 200 pontos distintos, distribuídos de maneira uniforme dentro da mesma. Essa escolha foi feita para se garantir uma grande quantidade de pontos de corte testados, além de garantir que os pontos gerados apresentassem a mesma precisão em casas decimais que os valores de  $\Delta$ .

No terceiro passo, para cada um dos pontos definidos na etapa anterior, são calculadas a Sensibilidade, a Especificidade, a Acuidade e o valor SP de todos os pontos de corte. Por fim, escolhe-se o ponto de corte dentre todos, que possua a maior medida SP. Abaixo, a listagem dos passos citados:

1. Determinar a faixa dinâmica de  $(\Delta_{Bov})$  ou  $(\Delta_{Bov} - \Delta_{Av})$
2. Dividir a faixa dinâmica uniformemente em 200 pontos
3. Para cada ponto, calcular: Sensibilidade, Especificidade, Acuidade e valor SP
4. Escolher dentre todos os cortes, o que com maior SP.

#### 4.4.2 Análise

O experimento foi feito para todas as fazendas disponíveis, sendo calculados pontos de corte baseados no valor SP para as fazendas UFF, Macaé e RS. Um quarto cenário ainda considerado no experimento é o caso em que os animais de UFF, Macaé e RS são colocados como pertencente a um único grupo e tratados juntos. Tal conjunto de dados é considerado como uma fazenda e é nomeada neste experimento como Fazenda *All*. O intuito para o uso da fazenda fictícia *All* é o fato desta representar uma abordagem global aos testes cervicais, apresentando-se como um cenário mais próximo ao de todos os rebanhos do Brasil.

Para todas as quatro fazendas consideradas, são analisados e comparados os testes intradérmicas pelo novo ponto de corte com a norma oficial brasileira. As medidas de performance consideradas nas análises são a Sensibilidade, a Especificidade, a Acuidade e o valor SP. Neste experimento, para a norma brasileira especificamente, um resultado dito inconclusivo é considerado como um diagnóstico errado pelo teste cervical.

A análise é feita separadamente para animais anérgicos e animais não anérgicos (TB+ e TB-). Apenas a fazenda RS não possui análise feita para bovinos anérgicos, já que este tipo de animal não está presente na fazenda.

# Capítulo 5

## Resultados

Neste capítulo, são apresentados os resultados referentes aos experimentos descritos no Capítulo 4. O capítulo segue a mesma ordem em que os experimentos foram apresentados.

A Seção 5.1 fala sobre o experimento sobre a adição de novos exames aos testes cervicais. Em seguida, são mencionados os resultados sobre análise de Repetibilidade da adição de exames complementares aos testes intradérmicos 5.2. Por último, são apresentados na Seção 5.3 os resultados sobre a otimização proposta aos pontos de corte definidos pela legislação brasileira.

### 5.1 Adição de exames

Nesta seção são abordados os resultados obtidos para o primeiro experimento proposto. O objetivo deste experimento foi avaliar o impacto que a adição de novos exames causa aos métodos de diagnósticos de tuberculose bovina hoje preconizados pela lei brasileira. Usou-se uma abordagem neural para avaliar tal impacto.

A seção apresenta os resultados separados por fazendas e em cada uma são apresentados os resultados dos modelos neurais e a eficiência dos exames usados. Uma última seção compara todas as fazendas e seus modelos neurais.

#### 5.1.1 Fazenda UFF

Para a fazenda UFF foram treinados 8 modelos neurais diferentes, divididos em dois grupos: os que receberam e os que não receberam animais anérgicos em seu treinamento. Os modelos neurais considerados em cada caso são:

- *TCS, TCC*
- *TCS, TCC, ELISA*
- *TCS, TCC, Leite*

- *TCS, TCC, ELISA, Leite*

## Modelos treinados sem animais anérgicos

Nesta modalidade de treinamento, os animais classificados como anérgicos não são incluídos nos conjuntos de treinamento dos modelos neurais. Foi usado como principal medida de análise dos modelos os seus valores SP, obtidos de cada uma das redes neurais geradas na validação cruzada. A Figura 5.1 apresenta o comportamento do valor SP para cada modelo neural, além de destacar sua média e desvio padrão. Observando os histogramas a partir do modelo *TCS TCC*, a medida que cada novo exame foi adicionado, ocorreu um deslocamento dos valores SP em direção à direita. Tal fato é visto pelo aumento das médias em cada um dos modelos, começando em 0,90 em *TCS TCC*, aumentando para 0,93 em *TCS TCC ELISA*, com um segundo aumento para 0,97 em *TCS TCC Leite* e por último, o modelo neural com todos os exames apresentando 0,99 de valor médio de SP.

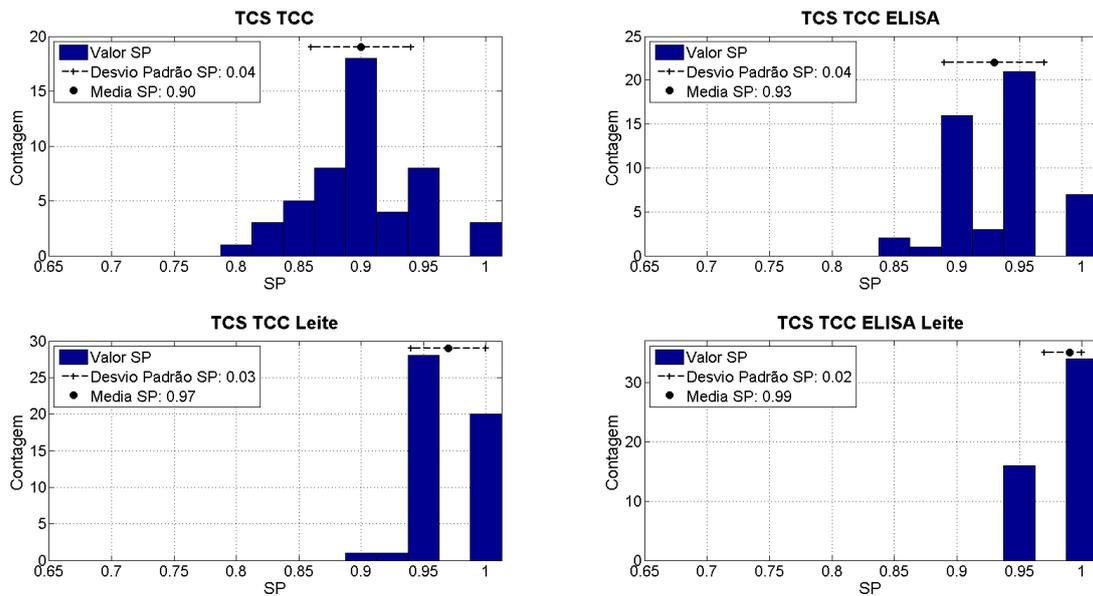
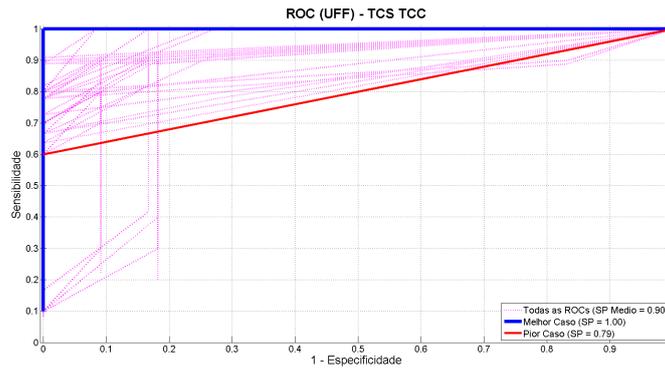
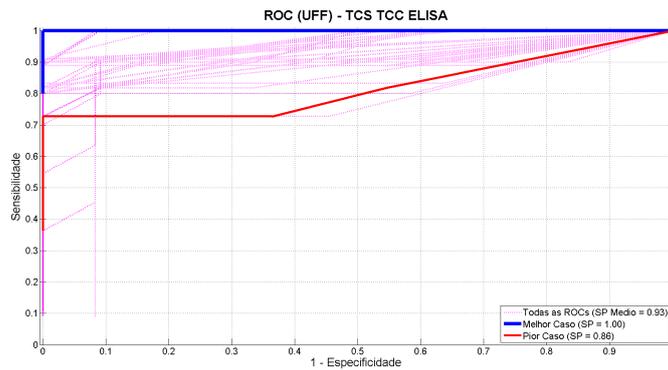


Figura 5.1: Distribuição do SP para os modelos da fazenda UFF **SEM** animais anérgicos no treinamento

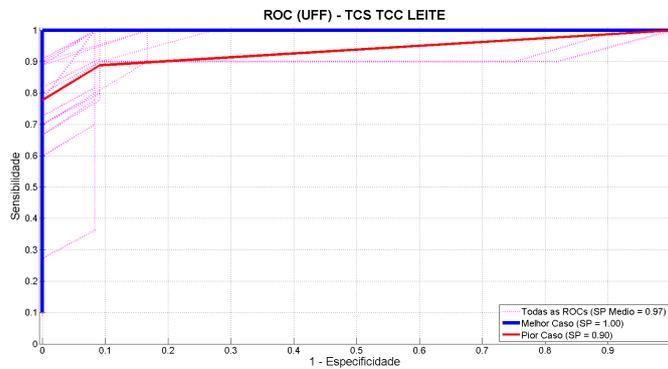
A análise por curva ROC (Figura 5.2) é análoga à feita para os valores SP, onde para cada modelo existem um total de 50 curvas diferentes e cada uma está associada a uma rede neural gerada na validação cruzada. A figura ainda destaca dentre todas as curvas, as de maior e menor SP. A avaliação de cada curva foi feita considerando-se a área abaixo da mesma e pode-se observar um aumento gradativo destas áreas à medida que cada novo exame é adicionado ao modelo neural anterior. Verificou-se que todas as curvas ROC, com exceção das que apresentaram o maior SP, sofreram um deslocamento para cima.



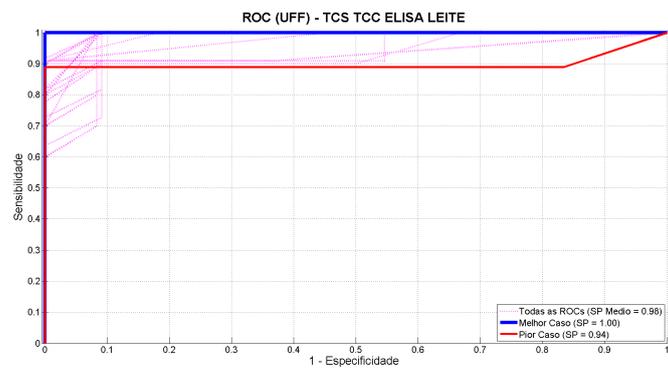
(a) Modelo TCS TCC



(b) Modelo TCS TCC ELISA



(c) Modelo TCS TCC Leite



(d) Modelo TCS TCC ELISA Leite

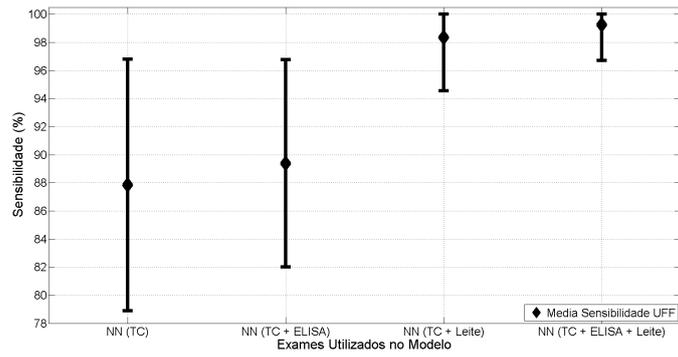
Figura 5.2: Gráficos de curva ROC para a fazenda UFF, SEM anérgicos no treinamento

A partir de cada curva ROC e da distribuição dos animais TB+ e TB-, é possível determinar as medidas de performance dos modelos neurais, sendo estas medidas a Sensibilidade, a Especificidade, a Acuidade e o SP. Cada medida se refere às redes neurais geradas na validação cruzada e a Figura 5.3 apresenta graficamente a média e desvio padrão para cada uma, considerando apenas os animais não anérgicos.

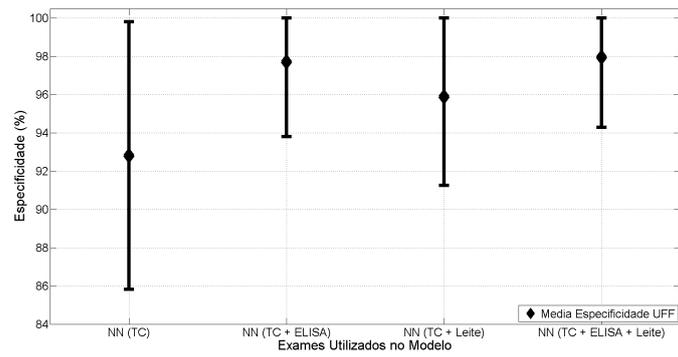
Para a medida de Sensibilidade, a Figura 5.3(a) mostra que o modelo possuindo apenas os testes cervicais apresentou um valor médio igual a 88% e com um desvio padrão de 7%. Quando o ELISA é adicionado, notou-se pouca mudança na medida, com a média estando em 89% e o seu desvio padrão em 7%, o que apesar de ser um ganho é visto como insuficiente para diferenciar os dois modelos e por isso, os modelos *TCS TCC* e *TCS TCC ELISA* são considerados estatisticamente equivalentes. Ao adicionar o leite aos testes cervicais, viu-se um grande aumento na média de Sensibilidade e uma diminuição de seu desvio padrão, respectivamente 98% e 4%, indicando que os modelos classificam melhor os animais TB+ e com uma menor variação em seu diagnóstico. Por fim, ao usar todos os exames disponíveis, a média e o desvio padrão foram respectivamente 99% e 2%, não indicando assim uma mudança grande do que foi visto apenas com o uso do leite.

Para a medida de Especificidade, a Figura 5.3(b) mostra o modelo *TCS TCC* com valor médio igual a 93%, com desvio padrão em 7%, porém, a adição do ELISA gera grande mudança na medida, aumentando a média até 98% com um desvio de 4%. Ao adicionar o exame de leite aos testes intradérmicos, a média encontrada é de 96% com desvio padrão de 5%, mostrando que há um aumento e um ganho na medida quando comparado apenas com os testes cervicais, porém com uma piora em comparação apenas com o ELISA. Quando os quatro exames são utilizados no modelo neural, o que se viu foi um desempenho idêntico para o modelo com o ELISA, indicando que usado em conjunto com o leite o modelo não perde em Especificidade por causa do ELISA.

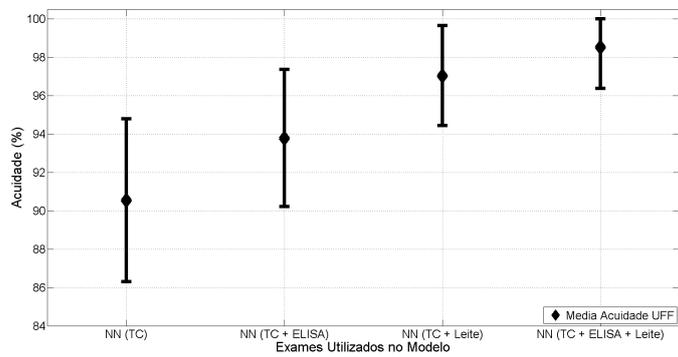
A Acuidade, relacionada na Figura 5.3(c), é dependente tanto da Sensibilidade quanto da Especificidade e por isso acompanhou os efeitos vistos nestas medidas. No modelo apenas com os testes cervicais, sua média e desvio padrão foi de, respectivamente, 91% e 4%, tendo aumentado para 94% e 4% ao adicionar o ELISA. Com a adição de leite, a média apresentou um ganho ainda maior indo até 97% e seu desvio padrão diminuindo para 3%. Com todos os exames sendo utilizados, a média e o desvio padrão encontrados foi de 99% e 2%. O aumento nesta medida quando o ELISA foi adicionado se deu principalmente pela Especificidade do exame, enquanto que no caso do leite, tal aumento ocorreu por causa de seu ganho de Sensibilidade.



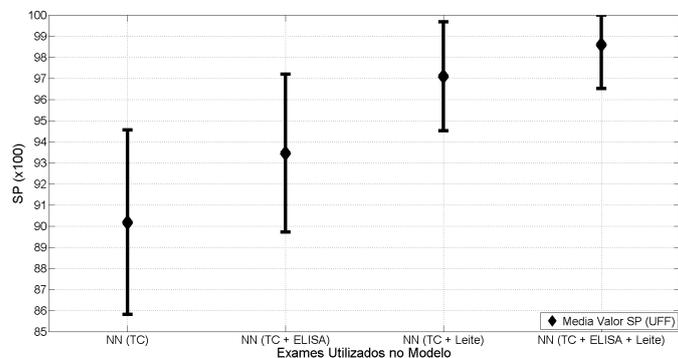
(a) Sensibilidade



(b) Especificidade



(c) Acuidade



(d) Valor SP

Figura 5.3: Média e Desvio Padrão das medidas de performance para a fazenda UFF, SEM anérgicos no treinamento

Para a medida SP (Figura 5.3(d)), o comportamento foi o mesmo encontrado para a Acuidade, com seus valores estando no formato decimal e não mais em valores de porcentagem. A Tabela 5.1 relaciona as medidas de desempenho para cada um dos modelos neurais da fazenda UFF, vistas na Figura 5.3.

Tabela 5.1: Performance das Redes neurais de UFF para animais não anérgicos. Treinamento SEM anérgicos

Mod. Neural	Sens. (%)	Espec. (%)	Acc. (%)	SP
TCS TCC	$88 \pm 9$	$93 \pm 7$	$91 \pm 4$	$0,90 \pm 0,04$
TCS TCC ELISA	$89 \pm 7$	[94 ; 100]	$94 \pm 4$	$0,93 \pm 0,04$
TCS TCC Leite	[94 ; 100]	[91 ; 100]	$97 \pm 3$	$0,97 \pm 0,03$
TCS TCC ELISA Leite	[96 ; 100]	[94 ; 100]	[97 ; 100]	[0,97 ; 1]

Um fator impactante na inclusão de novos exames aos testes cervicais, foi o fato de que qualquer um dos exames complementares adicionados gerou aumentos no valor médio e manteve ou diminuiu o desvio padrão de cada uma das medidas de performance nos modelos. Tal fato mostra que em UFF, tanto o ELISA quando o exame de leite ao serem adicionados tornam o classificador mais eficiente e aumentaram suas taxas de verdadeiro positivo e verdadeiro negativo. Dentre todos os modelos neurais testados, o modelo *TCS TCC ELISA Leite* apresentou as melhores medidas de performance, com as maiores médias e menores desvios padrão.

Em seguida, os modelos neurais foram testados para os animais anérgicos da fazenda e nesta análise foi considerada apenas a medida de performance de Sensibilidade. As demais medidas não puderam ser calculadas pois dependem de animais do tipo TB− e os anérgicos são sempre animais TB+. A Figura 5.4 apresenta a média e o desvio padrão da Sensibilidade de cada modelo neural para os animais anérgicos. O modelo neural apenas com os testes cervicais apresentou 84% de Sensibilidade e a faixa dinâmica desta medida variou de 64% até 100%. Ao adicionar o ELISA ao modelo neural, sua Sensibilidade média sobe para para 92% como também foi visto um pequeno aumento em seu desvio padrão, agora de 21% porém ainda muito próximo ao modelo *TCS TCC*. O terceiro modelo testado foi o modelo com o exame leite adicionado aos testes cervicais e neste modelo foi observado uma queda no valor médio da Sensibilidade e um aumento de seu desvio padrão para 28%, indicando uma maior variação da performance de classificação dos anérgico nas redes da validação cruzada. O último modelo neural testado, *TCS TCC ELISA Leite*, apresentou uma média igual a 80% e um desvio padrão de 24%. A Tabela 5.2 trás a relação dos valores de Sensibilidade dos modelos neurais para os animais anérgicos.

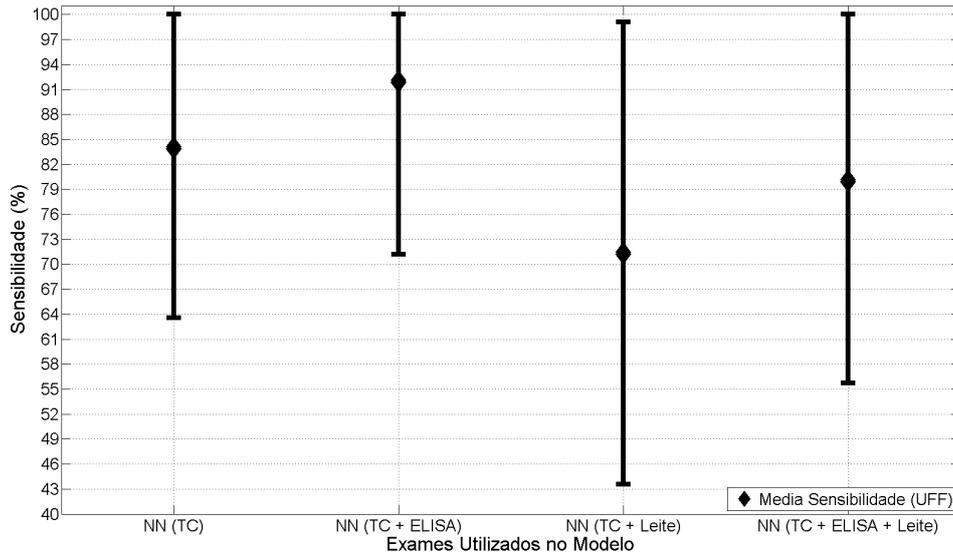


Figura 5.4: Sensibilidade dos animais anérgicos para os modelos da fazenda UFF SEM animais anérgicos no treinamento

Tabela 5.2: Performance das Redes neurais de UFF para animais anérgicos. Treinamento SEM anérgicos

Mod. Neural	Sensibilidade (%)
TCS TCC	[64 ; 100]
TCS TCC ELISA	[71 ; 100]
TCS TCC Leite	71 ± 28
TCS TCC ELISA Leite	[56 ; 100]

Dentre os quatro modelos neurais testados, o modelo *TCS TCC ELISA* foi o que apresentou a melhor Sensibilidade e tal resultado reforça o grande uso do ELISA para identificar anérgicos [50][51]. O exame leite ao ser adicionado aos exames cervicais gera uma queda na Sensibilidade, além de ser o modelo que possui a maior dispersão desta medida, indicando que tal exame não é apropriado para o diagnóstico de anérgicos.

### Modelos treinados com animais anérgicos

Nesta modalidade de treinamento, os animais classificados como anérgicos estão incluídos nos conjuntos de treinamento dos modelos neurais. Pelos gráficos gerados para os valores SP de cada um dos modelos neurais analisados, exibidos na Figura 5.5, a tendência observada para o treinamento sem os animais anérgicos se mantém para o treinamento atual. A cada novo exame adicionado aos exames cervicais há um aumento nos valores médios de SP em cada modelo neural, destacados nos

histogramas juntamente com os valores de desvios padrão, e que são percebidos pelo deslocamento à direita dos histogramas. A partir do modelo *TCS TCC*, o valor médio de SP visto para este modelo é de 0,90 e ao acrescentar o ELISA ao modelo, este valor sobe para 0,93. Com o exame de leite adicionado aos testes cervicais, o média de SP deste modelo é de 0,96 e finalmente, ao se usar, todos os exames foi obtido um valor médio de 0,98. Este aumento de valor médio é um fator que indica o ganho de classificação que os modelos obtiveram com a inclusão de exames complementares. Os histogramas ainda destacam os intervalos para o desvio padrão de SP em cada modelo, e o que se vê é que a cada novo exame adicionado aos testes cervicais, há uma diminuição dos desvios.

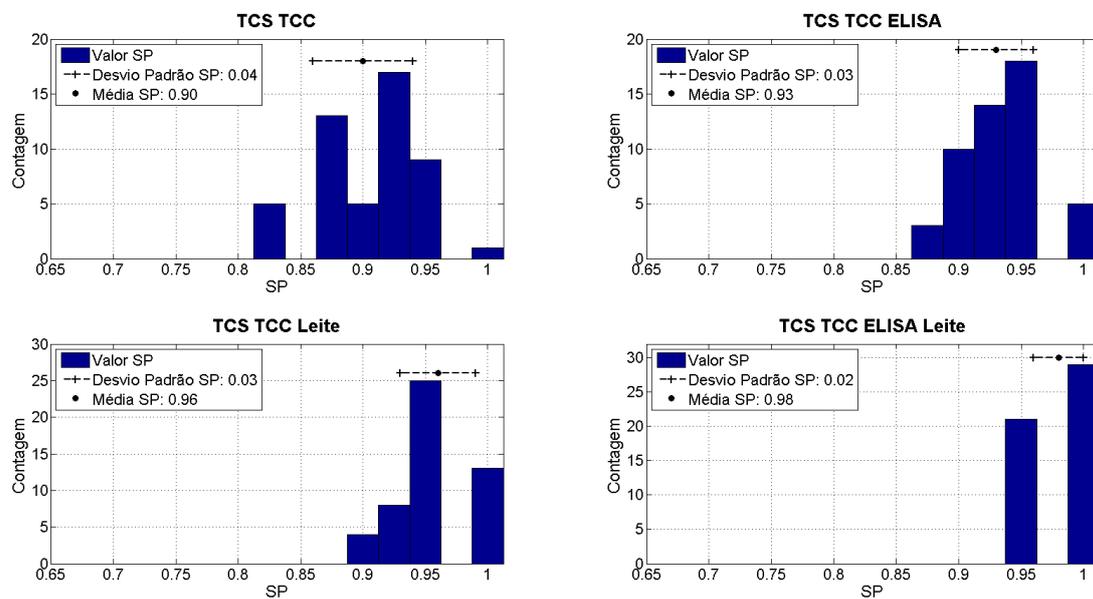
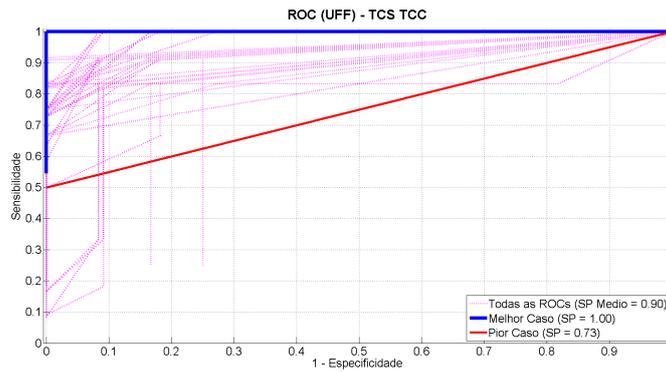
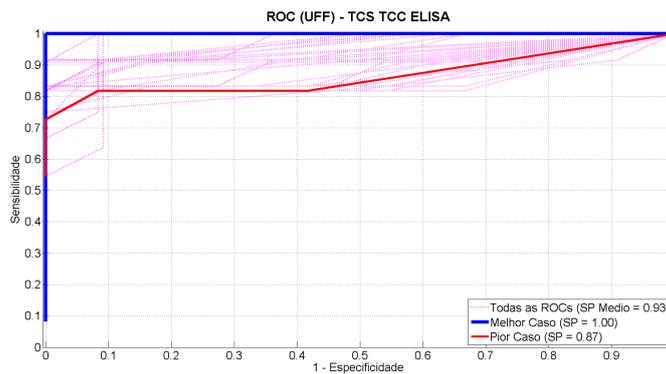


Figura 5.5: Distribuição do SP para os modelos da fazenda UFF **COM** animais anérgicos no treinamento

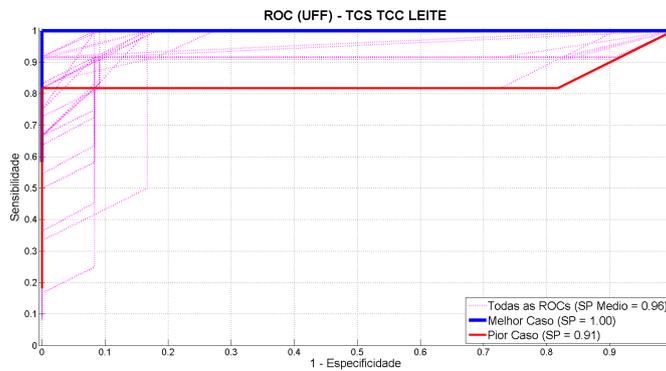
As curvas ROC geradas para cada um dos modelos neurais (Figura 5.6) são usadas primeiramente para a determinação das médias de Sensibilidade, Especificidade, Acuidade e mesmo SP dos modelos neurais. Cada modelo possui um conjunto de curvas ROC e a partir de todas as curvas, pode-se calcular média e desvio padrão de cada medida. Uma análise também é feito por meio das curvas ROC, tomando-se como parâmetro a área abaixo de cada curva para avaliar cada modelo neural. As curvas destacadas para o valor máximo SP de cada rede neural apresentaram o mesmo desempenho, com uma classificação perfeita dos animais apresentados em seus grupos de teste, mas para as demais curvas houve variação e conseqüentemente, diferentes áreas sob as mesmas. Para cada novo exame adicionado aos testes cervicais, notou-se um aumento da área sob as curvas, com o exame leite apresentando maiores áreas do que o ELISA e ainda maiores áreas para os modelos treinados com o ELISA e o exame de leite juntos.



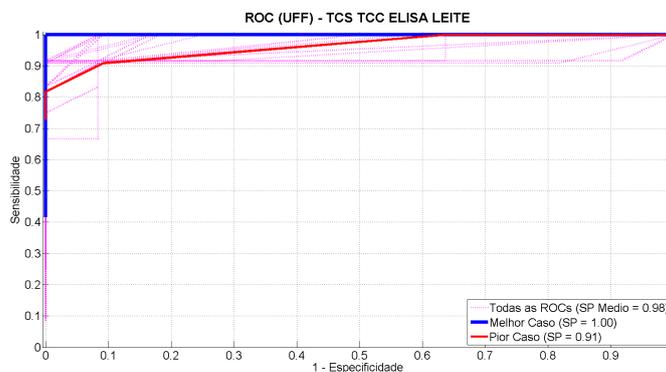
(a) Modelo TCS TCC



(b) Modelo TCS TCC ELISA



(c) Modelo TCS TCC Leite



(d) Modelo TCS TCC ELISA Leite

Figura 5.6: Gráficos de Curva ROC para a fazenda UFF, COM Anérgicos no Treinamento

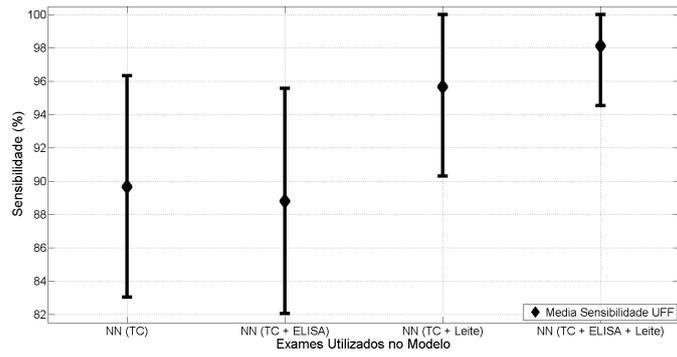
A Figura 5.7 relaciona graficamente as medidas de performance dos modelos neurais estudados, apresentando a média e seus desvios padrão. A figura relaciona apenas os animais do tipo não anérgicos.

Para a Sensibilidade, a Figura 5.7(a) mostra que o modelo *TCS TCC* apresentou um valor médio igual à 90% e com um desvio padrão de 7%. Ao adicionar ELISA, notou-se novamente pouca mudança na medida, com a média diminuindo para 89% e mantendo seu desvio padrão em 7%, indicando que apesar da queda do valor médio, os modelos *TCS TCC* e *TCS TCC ELISA* podem ser considerados estatisticamente equivalentes e assim a adição de ELISA não impacta na Sensibilidade dos modelos. Ao adicionar o leite aos testes cervicais, a figura mostra um grande aumento na média de Sensibilidade, para 96%, e uma diminuição de seu desvio padrão para 5%, indicando que os modelos classificam melhor os animais TB+ e possui uma menor variação em seu diagnóstico. Por fim, ao se usar todos os exames disponíveis, a média e o desvio padrão foram respectivamente 98% e 4%, mostrando que este modelo neural possui a melhor performance de classificação de animais TB+.

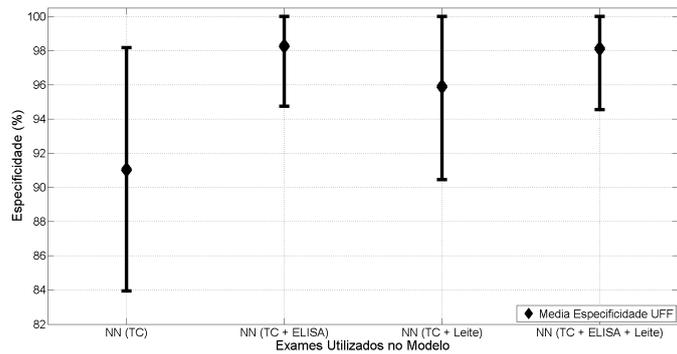
A Especificidade dos modelos, mostrado na Figura 5.7(b) relaciona o modelo *TCS TCC* com valor médio igual à 91% e desvio padrão em 7%. A adição de ELISA gera na medida um aumento para seu valor médio, indo até 98%, e a diminuição de seu desvio padrão, caindo para 4%. O uso do exame de leite junto aos testes cervicais aumenta a média, quando comparado apenas o modelo *TCS TCC* até de 96% e também diminui o desvio padrão até 5%, mostrando que há um ganho na medida com a adição do exame. No entanto, comparando o impacto do exame de leite com o visto para o ELISA, o teste sorológico apresenta ganhos um pouco maiores. Por fim, o modelo neural com os quatro exames juntos apresenta um desempenho idêntico ao modelo *TCS TCC ELISA*, indicando que o exame de leite ao ser usado com o ELISA não gera as perdas vistas quando usado isoladamente com os testes cervicais.

A Figura 5.7(c) apresenta a Acuidade dos modelos neurais, que segue o comportamento visto para a Sensibilidade e a Especificidade, obtendo melhores valores de performance à cada exame adicionado. Inicialmente, para os testes cervicais apenas, o valor médio e o desvio padrão são respectivamente 90% e 4%. Com a adição de ELISA ao modelo, os valores são de 94% e 3% para a média e o desvi padrão. Com o uso do exame leite, o desvio padrão é o mesmo de ELISA, com 3% porém, há um ganho maior na média, com 96%. O modelo neural com todos os exames apresentou um valor médio de 98% e um desvio padrão de 2%, indicando ganho na performance do modelo para esta medida.

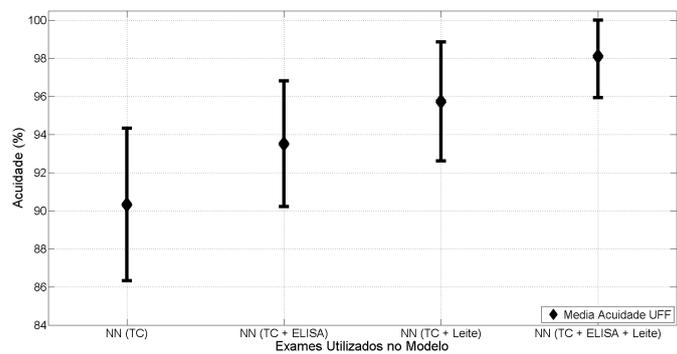
A Figura 5.7(d) relaciona a medida de SP e o seu comportamento foi o mesmo do encontrado para a Acuidade, com seus valores estando no formato decimal e não mais em valores de porcentagem.



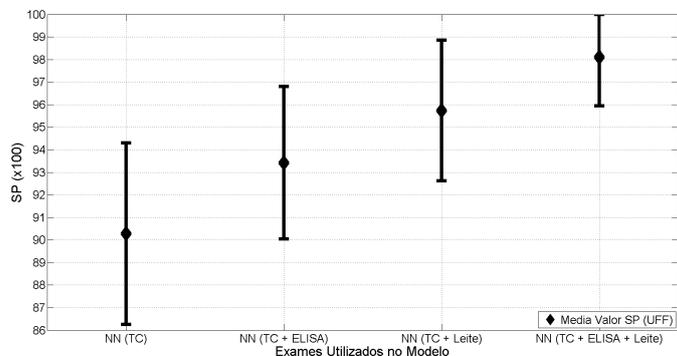
(a) Sensibilidade



(b) Especificidade



(c) Acuidade



(d) Valor SP

Figura 5.7: Média e Desvio Padrão das medidas de performance para a fazenda UFF, COM anérgicos no treinamento

A Tabela 5.3 apresenta as medidas de desempenho observadas na Figura 5.7. O que foi observado para os modelos treinados com anérgicos foi o mesmo para o treinamento sem estes animais, onde a adição de novos exames junto aos testes cervicais aumenta a eficiência na classificação destes animais. Novamente foi observado que a adição de ELISA aumenta a Especificidade dos modelos, enquanto o exame de leite apresenta ganhos na Sensibilidade. O maior ganho em Acuidade e SP foi obtido para o modelo *TCS TCC ELISA Leite*, mostrando que o uso dos dois exames complementares juntos retorna a maior eficiência na classificação dos bovinos, já que este modelo possui a maior média e também o menor desvio padrão para estas medidas.

Tabela 5.3: Performance das Redes neurais de UFF para animais não anérgicos. Treinamento **COM** anérgicos

<b>Mod. Neural</b>	<b>Sens. (%)</b>	<b>Espec. (%)</b>	<b>Acc. (%)</b>	<b>SP</b>
TCS TCC	$90 \pm 7$	$91 \pm 7$	$90 \pm 4$	$0,90 \pm 0,04$
TCS TCC ELISA	$89 \pm 7$	[94 ; 100]	$94 \pm 3$	$0,93 \pm 0,03$
TCS TCC Leite	[91 ; 100]	[91 ; 100]	$96 \pm 3$	$0,96 \pm 0,03$
TCS TCC ELISA Leite	[94 ; 100]	[94 ; 100]	$98 \pm 2$	$0,98 \pm 0,02$

Para os animais anérgicos a análise dos modelos neurais considera apenas a medida de Sensibilidade. A Figura 5.8 trás a relação de cada modelo neural com seus valores de média e desvio padrão. No modelo *TCS TCC* o valor médio de Sensibilidade foi de 93%, variando de 79% até 100% na classificação destes animais. O modelo neural com o ELISA apresenta um aumento em sua média, comparado apenas ao modelo dos testes cervicais, até 97% e também apresenta uma diminuição de desvio padrão para 11% e variando sua Sensibilidade de 86% até 100%, indicando que o modelo neural com ELISA possui menor incerteza na classificação dos anérgicos. Com o exame de leite, novamente se observa uma queda em seu valor médio e também um aumento no desvio padrão do mesmo para 20%. O modelo neural usando todos os exames novamente apresenta performance muito próxima ao modelo *TCS TCC ELISA*, porém com uma média um pouco menor e um desvio padrão um pouco maior, respectivamente 95% e 12%. A Tabela 5.4 contém os valores observados na Figura 5.8 para a performance dos modelos sobre os anérgicos apenas.

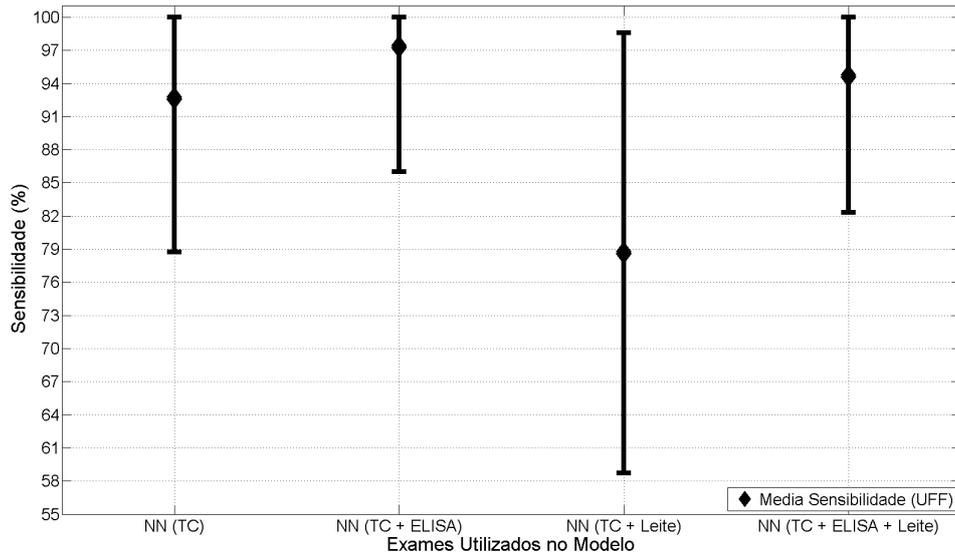


Figura 5.8: Sensibilidade dos animais anérgicos para os modelos da fazenda UFF COM animais anérgicos no treinamento

Tabela 5.4: Performance das Redes neurais de UFF para animais anérgicos. Treinamento COM anérgicos

Mod. Neural	Sensibilidade (%)
TCS TCC	[79 ; 100]
TCS TCC ELISA	[86 ; 100]
TCS TCC Leite	79 ± 20
TCS TCC ELISA Leite	[83 ; 100]

Como visto no caso dos modelos treinados sem animais anérgicos, o modelo neural usando apenas ELISA junto aos testes cervicais se mostrou mais eficaz na identificação destes animais, enquanto o exame de leite apresentou uma diminuição na medida de Sensibilidade do modelo. Notou-se ainda que o treinamento com os anérgicos gerou um aumento da Sensibilidade média e uma diminuição no desvio padrão para todos os modelos neurais desenvolvidos quando comparados ao treinamento desconsiderando estes bovinos (Tabelas 5.2 e 5.4). Este fato mostrou que o uso destes animais no treinamento foi benéfico para a classificação dos mesmos pelos modelos neurais.

### Exames Isolados

A análise por exames isolados em UFF é feita de modo a considerar os resultados obtidos por cada exame sozinho e isolado sobre todos os animais da fazenda. Para

cada exame, é calculado as medidas de Sensibilidade, Especificidade, Acuidade e seu valor SP através de seus resultados disponibilizados nos dados da fazenda.

A Tabela 5.5 exibe a relação dos resultados relativos à cada exame somente sobre os animais não anérgicos. Os exames isolados de TCS e TCC são os que apresentam os maiores valores de SP, respectivamente 0,78 e 0,82. Para o exame TCC, sua Especificidade atinge 100%, contra 87% do TCS. Para a Sensibilidade, o TCS apresentou maior eficiência, com 70% para esta medida, contra 65% do TCC. A Acuidade de ambos foi de 79% e 84% para, respectivamente, o TCS e o TCC. O ELISA não apresentou bons valores de performance, com 45% de Sensibilidade, 57% de Especificidade, 51% de Acuidade e 0,51 de valor SP. Por fim, o exame de leite apresentou medidas muito baixas de performance, com 15% de Sensibilidade, 4% de Especificidade, 9% de Acuidade e 0,09 de valor SP.

Tabela 5.5: Performance dos exames isolados de UFF para animais não anérgicos

<b>Exame</b>	<b>Sens. (%)</b>	<b>Espec. (%)</b>	<b>Acc. (%)</b>	<b>SP</b>
TCS	70	87	79	0,78
TCC	65	100	84	0,82
ELISA	45	57	51	0,51
Leite	15	4	9	0,09

Os resultados tanto dos modelos neurais usando apenas os testes cervicais como também seus exames isolados confirmam que tanto o TCS quanto o TCC não são capazes de encontrar corretamente todos os animais com Tuberculose. O ELISA apresentou resultados que mostram que este exame também não é capaz sozinho identificar todos os bovinos, porém pode ser usado de maneira complementar ao diagnóstico da TB. Uma característica importante vista é a sua maior Especificidade em animais não anérgicos, comparando com sua Sensibilidade. Tal medida explica os aumentos de Especificidade vistos nos modelos neurais quando ELISA foi acrescentado aos modelos (Figuras 5.3 e 5.7). O maior valor de Especificidade por parte de ELISA é visto também em [15], onde os resultados mostraram que ELISA apresentou sempre uma maior Especificidade nos bovinos testados, mesmo em diferentes momentos em que o exame foi aplicado.

O exame leite, quando comparado aos outros exames isolados, apresentou valores baixos em todas as medidas de performance consideradas. Tal resultado pode ser explicado pelo fato do exame leite não ter sido aplicado em todos os animais da fazenda UFF. Dos 46 animais de UFF, apenas 16 animais (35% do total) tiveram o exame aplicado, com os 30 animais restantes (65% do total), incluídos os animais anérgicos de UFF, considerados com resultados inconclusivos. Ao se considerar apenas os 16

bovinos em que o exame foi aplicado, 14 (88%) eram animais tuberculosos e apenas 2 (12%) eram animais livres de *M. bovis*. Dentre os 14 bovinos TB+, o exame de leite detectou o agente *M. bovis* em 4 animais, obtendo assim uma Sensibilidade real de 29%, enquanto que para os dois animais TB– o exame classificou corretamente 1 bovino, obtendo uma Especificidade real de 50%. Um fator importante visto para o exame de leite foi que todos os animais TB+ foram classificados como “Negativo” ou “Inconclusivo” pelos testes cervicais, ou seja, o exame de leite conseguiu identificar animais doentes que os testes oficiais não conseguiram, o que explica o aumento de Sensibilidade dos modelos neurais quando o leite é adicionado aos testes cervicais. A Tabela 5.6 relaciona as medidas de performance real do exame de leite, ou seja, apenas para os animais não anérgicos que foram aplicados o exame. A tabela apresenta a quantidade de animais por diagnóstico, e a quantidade de animais que o exame conseguiu classificar corretamente e medida de performance relacionada a cada um dos tipos de animais relacionados.

Tabela 5.6: Performance de identificação de animais não anérgicos testados com o exame de leite

<b>Diagnóstico</b>	<b>Ocorrências</b>	<b>Classificação Correta</b>	<b>Medida</b>
TB+	14 (88%)	4 (29%)	Sensibilidade
TB–	2 (12%)	1 (50%)	Especificidade
<b>Total</b>	<b>16 (100%)</b>	<b>5 (31%)</b>	Acuidade

Para os animais anérgicos de UFF, a Tabela 5.7 relaciona as Sensibilidades obtidas por cada exame isoladamente sobre estes bovinos. Para os testes cervicais, as Sensibilidades foram de 33% (1 animal) e 0%, respectivamente para o TCS e o TCC. O exame de leite foi outro exame que apresentou 0% na identificação dos anérgicos e por fim, o ELISA foi o exame que apresentou 100% na Sensibilidade.

Tabela 5.7: Performance dos exames isolados de UFF para animais anérgicos

<b>Exame</b>	<b>Sensibilidade (%)</b>
TCS	33
TCC	0
ELISA	100
Leite	0

Pelo conceito de anergia, exposto neste estudo na Seção 2.2.6, um animal anérgico é assim considerado quando este, mesmo estando com tuberculose, não reage aos

testes intradérmicos nele aplicados. Para a fazenda UFF, estes animais foram assim definidos ao não serem identificados pelo TCC apenas. Tal fato explica o porque do TCC apresentar 0% de eficiência. O TCS também não é um exame que não é apropriado para a identificação de anérgicos, porém é um teste com maior Sensibilidade que o TCC, o que acarreta em maior classificação de animais TB+. A Sensibilidade de 33% vista para o teste se deu por este conseguir classificar 1 dos animais, dentre os 3 considerados. Os resultados e a discussão sobre os testes cervicais, inclusive para os animais anérgicos, são apresentados na Seção 5.3.

O ELISA foi o teste que obteve 100% na identificação dos anérgicos e é um teste eficaz para a identificação deste tipo de animal [47]. Os resultados para o exame isolado também confirmam os resultados obtidos com os modelos neurais desenvolvidos, onde modelos com a presença de ELISA apresentou aumento na taxa de Sensibilidade destes animais e assim corroborando o seu uso para anergia em bovinos.

O exame do leite, apesar de não ter conseguido identificar nenhum animal, não foi obtida nenhuma conclusão a respeito do exame aplicado à animais anérgicos. O exame apresentou uma limitação em sua aplicação, como visto para os animais não anérgicos, que foi limitada à uma parcela pequena da fazenda e dentre os animais testados por este exame, nenhum dos anérgicos foi testado. Logo, a razão para os resultados obtidos tanto pelos modelos neurais quanto para o exame de leite isolado se deu pela não utilização deste exame nestes animais.

O experimento feito em UFF mostrou que a abordagem multi-diagnóstica para a Tuberculose Bovina apresenta benefícios, como o aumento da Especificidade e da Sensibilidade em animais não anérgicos, respectivamente, devido à ELISA e ao exame de leite, como também para identificação de anérgicos pelo ELISA. Os modelos neurais vistos para esta fazenda sugerem que o uso combinado dos testes cervicais, de ELISA e de exames de leite apresenta a melhor eficiência para o diagnóstico da doença.

### 5.1.2 Fazenda Macaé

Para a fazenda Macaé, foram treinados 4 modelos neurais diferentes. Os modelos foram divididos, assim como na fazenda UFF, em cenários onde estes foram treinados na presença e na ausência de animais anérgicos. Como a fazenda Macaé possui apenas os exames cervicais e o ELISA, os modelos neurais desenvolvidos foram:

- *TCS, TCC*
- *TCS, TCC, ELISA*

## Modelos treinados sem animais anérgicos

Nesta modalidade de treinamento os animais classificados como anérgicos não são incluídos nos conjuntos de treinamento dos modelos neurais. Considerando a medida de performance SP, a Figura 5.9 trás sua distribuição para cada um dos modelos neurais desenvolvidos na fazenda Macaé. Cada histograma representa um modelo neural, onde seus valores de SP foram retirados de suas redes neurais geradas na validação cruzada. O primeiro fator observado é a semelhança entre as distribuições de SP em ambos os modelos que é refletido nos valores médios e em seus desvios padrão, destacados na figura. Em ambos os modelos, um desvio padrão de 0,05 foi encontrado, enquanto que suas médias de SP foram 0,86 para o modelo *TCS TCC* e 0,87 para o modelo *TCS TCC ELISA*. Tais valores indicam que os modelos neurais possuem performance muito similar, ou seja, que para esta medida a adição de ELISA não apresentou efeitos significativos para o modelo.

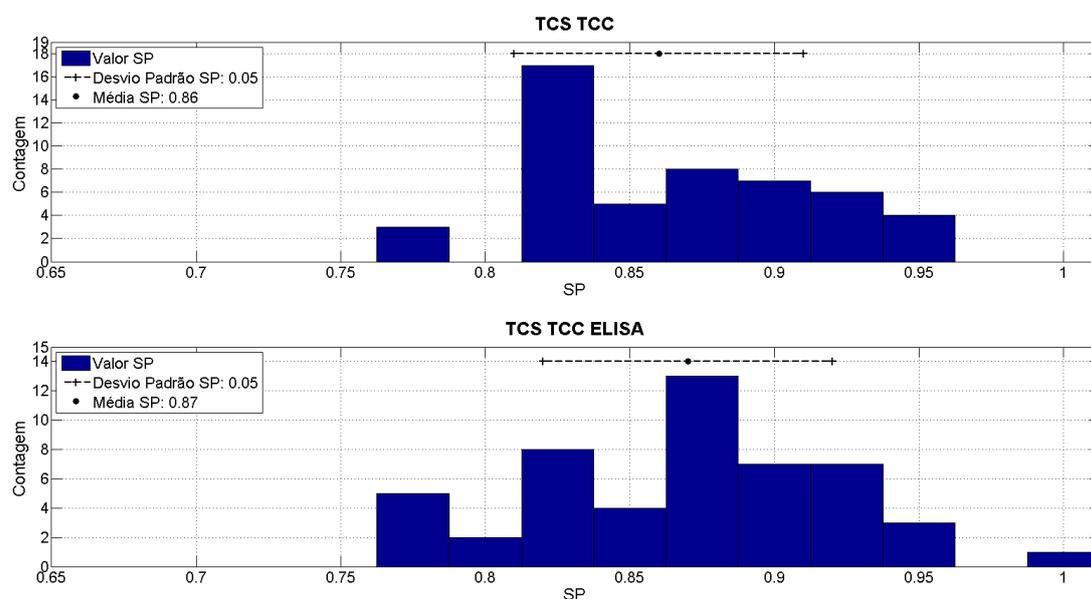
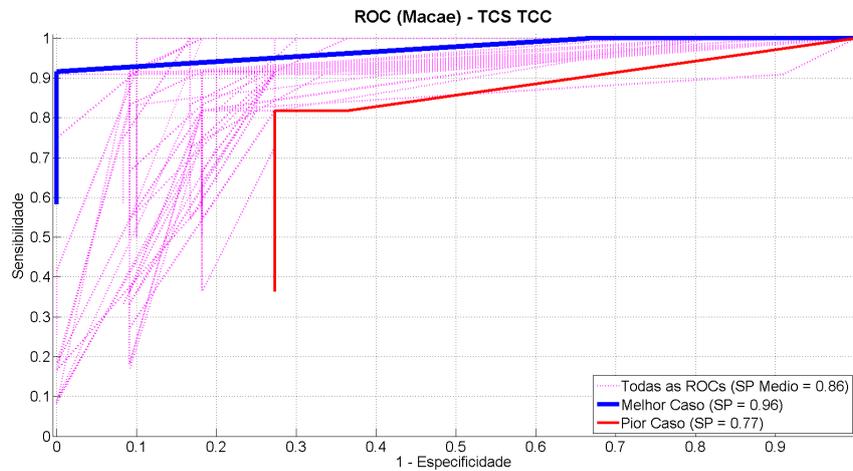


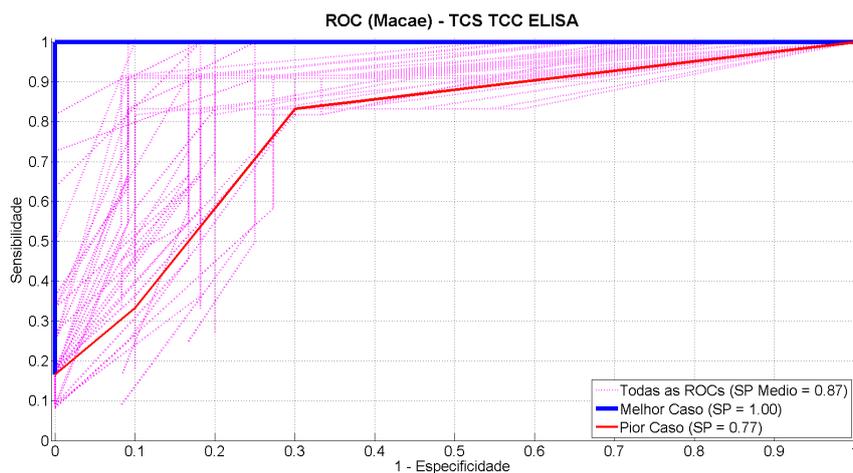
Figura 5.9: Distribuição do SP para os modelos da fazenda Macaé **SEM** animais anérgicos no treinamento

A Figura 5.10 apresenta as curvas ROC geradas para todas as redes neurais dos modelos considerados, tal como a indicação das curvas com maior e menor SP associado. Comparando apenas as curvas de pior SP, não notou-se graficamente grande diferença entre ambas, o que é também explicado por seus valores idênticos de SP (0,77). Para as curvas de maior SP, o modelo neural *TCS TCC ELISA* apresenta uma ROC perfeita, corroborado por seu valor SP igual a 1, enquanto o modelo *TCS TCC* apresentou 0,96 de medida SP para sua melhor performance. Para as demais curvas, ao considerar a métrica de área abaixo das mesmas, novamente não foi notado grande diferença entre todas as ROCs dos modelos neurais. Apesar das

curvas de maior e menor SP dos dois modelos serem destacadamente diferentes, as demais se mostram similares e isto é confirmado pelas médias de SP muito próximas (Figura 5.9). Com isso, a análise das curvas ROC por estes critérios não apontam ganho com a adição de ELISA aos testes cervicais.



(a) Modelo TCS TCC



(b) Modelo TCS TCC ELISA

Figura 5.10: Gráficos de Curva ROC para a fazenda Macaé, SEM Anérgicos no Treinamento

A partir das curvas ROC é ainda possível extrair as medidas de Sensibilidade, Especificidade e Acuidade para avaliar a performance dos modelos neurais. A Figura 5.11 apresenta graficamente, para cada medida de performance, o seu valor médio encontrado tal como seu desvio padrão, somente para os animais não anérgicos da fazenda.

Para a medida de Sensibilidade, a Figura 5.11(a) mostra um comportamento similar à medida em ambos modelos neurais. O modelo utilizando apenas os testes cervicais (*TCS TCC*) apresentou uma média igual a 90%, com seu desvio padrão

em 6%. Quando o ELISA é adicionado aos testes cervicais (modelo *TCS TCC ELISA*) o valor médio de Sensibilidade do modelo aumenta em 1%, passando para 91%, com seu desvio padrão mantendo-se o mesmo do modelo *TCS TCC*, com 6%. O mesmo valor de desvio padrão indica que os dois modelos possuem a mesma variação estatística de Sensibilidade e ao terem seus valores médios tão próximos, tais resultados mostram que os modelos são estatisticamente equivalentes. Esta equivalência é traduzida em uma igual performance na classificação de bovinos TB+ pelos dois modelos neurais.

Para a medida de Especificidade, exposta na Figura 5.11(b), os dois modelos neurais apresentam a mesma medida de Especificidade média, igual a 83%. Para o desvio padrão, o modelo *TCS TCC* apresentou uma dispersão igual a 7%, enquanto o modelo com ELISA adicionado o desvio padrão encontrado foi de 8%. Novamente, os valores próximos da média e do desvio padrão indicam que os modelos são similares para esta medida e que são estatisticamente equivalentes. Em Macaé não foi observado os ganhos de Especificidade proporcionados pela adição de ELISA vistos em UFF.

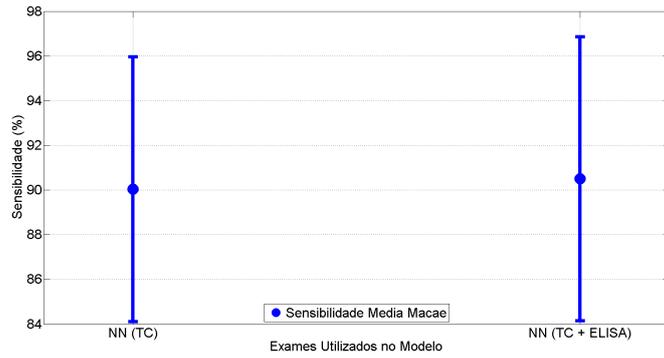
A Figura 5.11(c) apresenta a relação de resultados para a Acuidade dos modelos neurais. A Acuidade é uma medida diretamente relacionada às medidas de Sensibilidade e Especificidade, portanto seu comportamento foi similar ao visto em ambas. O modelo *TCS TCC* apresentou Acuidade média de 86% e um desvio padrão de 5%, enquanto o modelo *TCS TCC ELISA* apresentou média de 87%, apenas 1% maior que o anterior, com um desvio padrão de 5%. O mesmo desvio padrão indica que os dois modelos apresentaram variação similar em suas classificações, além da diferença na performance média ser muito pequena, fazendo com que os modelos fossem considerados equivalentes.

Para a medida de SP, apresentada na Figura 5.11(d), teve suas medidas também analisadas na Figura 5.9 e seus resultados mostram que os dois modelos apresentaram mesma performance na classificação dos bovinos não anérgicos de Macaé.

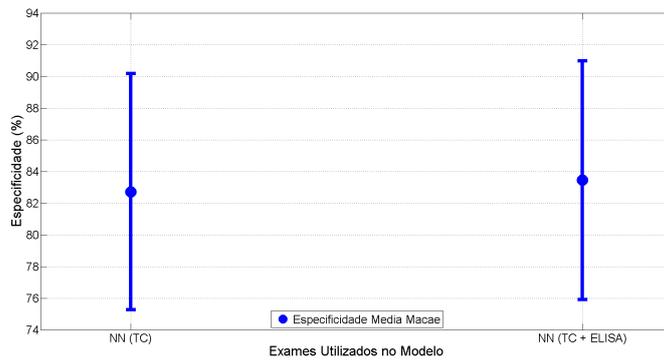
A Tabela 5.8 relaciona os valores encontrados para as medidas de performance vistas em 5.11.

Tabela 5.8: Performance das redes neurais de Macaé para animais não anérgicos. Treinamento **SEM** anérgicos

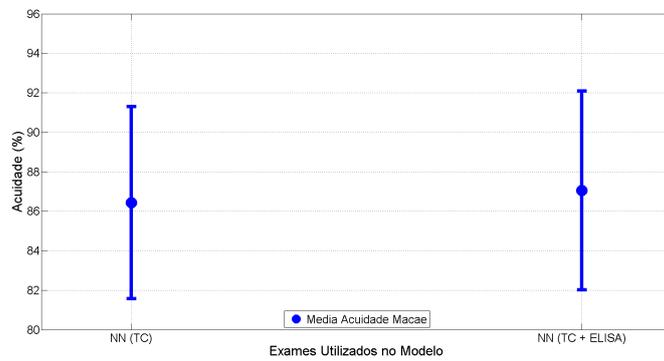
Mod. Neural	Sens. (%)	Espec. (%)	Acc. (%)	SP
TCS TCC	90 ± 6	83 ± 7	86 ± 5	0,86 ± 0,05
TCS TCC ELISA	91 ± 6	83 ± 8	87 ± 5	0,87 ± 0,05



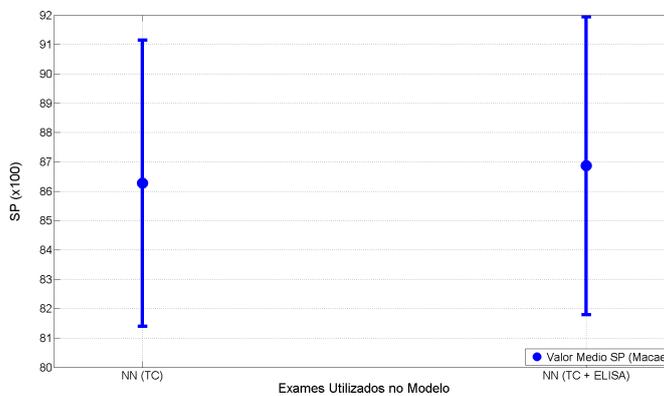
(a) Sensibilidade



(b) Especificidade



(c) Acuidade



(d) Valor SP

Figura 5.11: Média e Desvio Padrão das medidas de performance para a fazenda Macaé, SEM anérgicos no treinamento

Foi visto para os animais não anérgicos que o uso combinado dos testes cervicais com ELISA não apresentou benefícios para os modelos neurais treinados sem anérgicos, sem ganhos na Sensibilidade ou Especificidade.

A análise dos modelos neurais também foi feita para os animais anérgicos da fazenda, e por isso apenas a medida de Sensibilidade é considerada para estes animais. A Figura 5.12 relaciona graficamente a média e o desvio padrão da Sensibilidade de cada modelo. O modelo *TCS TCC* apresenta uma Sensibilidade média de 1% e um desvio padrão de 7%, que no entanto faz tal medida variar de 0 até 8%. O modelo com ELISA apresenta uma média de 11% e um desvio padrão de 23%, fazendo com que seu valor varie de 0% até no máximo 34%. A Tabela 5.9 relaciona os intervalos de variação de Sensibilidade dos modelos.

O modelo *TCS TCC* mostrou baixa performance para anérgicos e apresentou um desempenho médio próximo de 0%, indicando que o modelo não possui capacidade de classificação para estes animais. Ao adicionar ELISA, observou-se um aumento de 10 pontos percentuais na Sensibilidade média em relação ao modelo *TCS TCC*, o que ainda representa uma baixa classificação pois a fazenda possui 4 animais anérgicos e cada animal corretamente classificado representa aumento de 25% para esta medida em uma rede neural. No entanto, a adição do ELISA gera um impacto positivo na classificação, ainda que com um aumento da medida de dispersão. Tal aumento pode ser explicado pela eficiência do teste ELISA para a identificação de anérgicos.

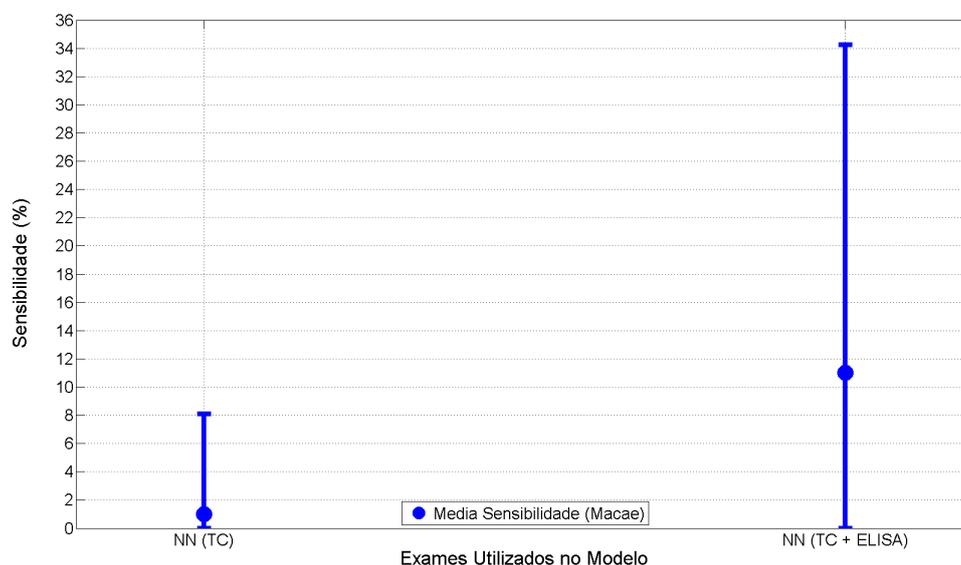


Figura 5.12: Sensibilidade dos animais anérgicos para os modelos da fazenda Macaé SEM animais anérgicos no treinamento

A Tabela 5.9 trás a relação da eficiência dos modelos neurais sobre os animais anérgicos.

Tabela 5.9: Performance das redes neurais de Macaé para animais anérgicos. Treinamento SEM anérgicos.

Mod. Neural	Sensibilidade (%)
TCS TCC	[0 ; 8]
TCS TCC ELISA	[0 ; 34]

No treinamento sem os bovinos anérgicos, o acréscimo de ELISA não apresentou ganhos na identificação da Tuberculose para os animais do tipo não anérgicos, enquanto para os anérgicos o acréscimo do exame juntos aos testes cervicais foi benéfico e apresentou um aumento de Sensibilidade, ainda que esta tenha se mantido com taxas baixas na classificação.

### Modelos treinados com animais anérgicos

Nesta modalidade de treinamento, os animais classificados como anérgicos estão incluídos nos conjuntos de treinamento dos modelos neurais.

A primeira análise para esta modalidade é feita sobre os valores SP dos modelos neurais. A Figura 5.13 apresenta histogramas com a distribuição das medidas SP, geradas pela validação cruzada, com a média e o desvio padrão de cada um destacados.

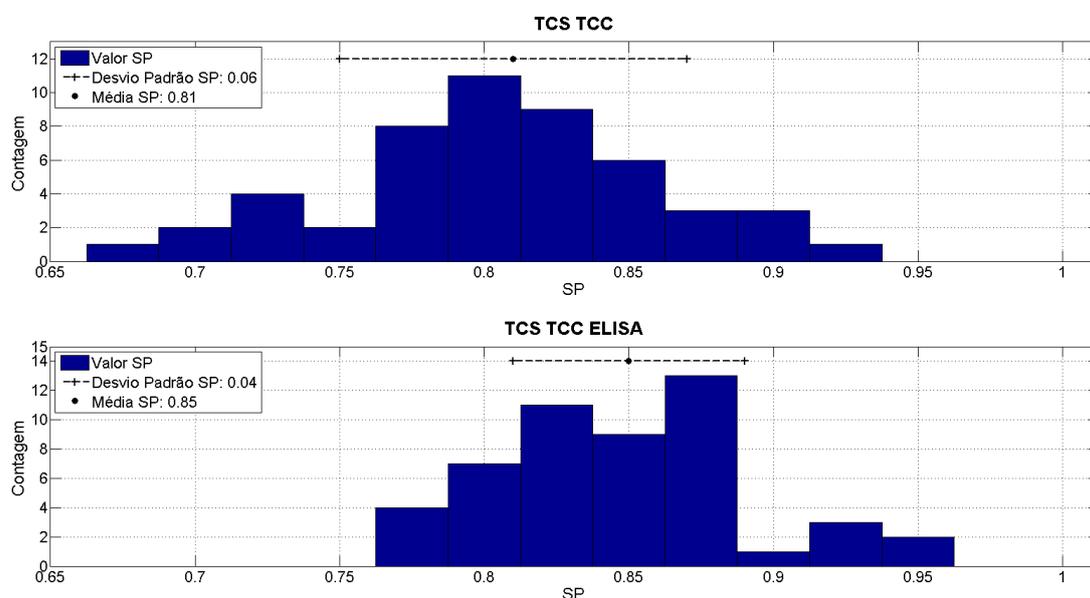
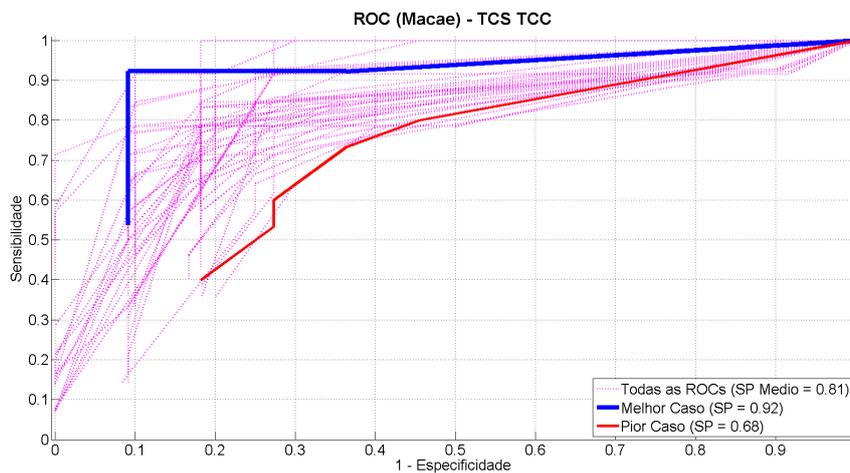


Figura 5.13: Distribuição do SP para os modelos da fazenda Macaé COM animais anérgicos no treinamento

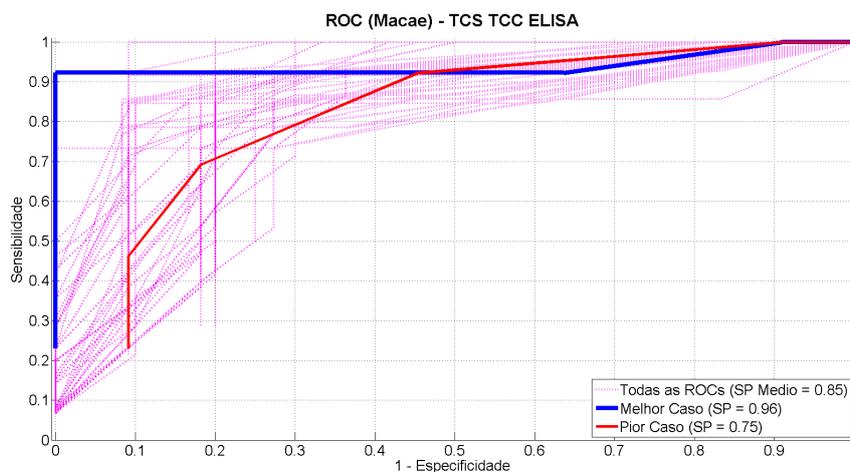
Analisando tal distribuição de SP notou-se primeiramente que o caso de treinamento com animais anérgicos gerou uma diminuição dos valores SP se comparados

com o treinamento sem animais anérgicos. Tal fato é visível pelo evidente deslocamento à esquerda pelas distribuições ao se comparar com a Figura 5.9. Ao compararmos apenas os dois modelos neurais, percebe-se um deslocamento à direita dos valores do modelo com ELISA contra o modelo apenas com os testes cervicais, indicando um aumento dos seus valores. Por meio dos valores médios e dos desvios padrão, tal deslocamento fica claro já que o modelo *TCS TCC* apresenta para o SP valor médio igual a 0,81 e 0,06 de desvio, enquanto o modelo *TCS TCC ELISA* apresenta para as mesmas medidas, respectivamente, 0,85 e 0,04. O aumento da média confirma o deslocamento das saídas dos modelos e a queda no desvio padrão indica uma menor dispersão do SP em torno de sua média.

Outra análise é feita através das curvas ROC pertencentes às redes neurais de cada um dos modelos, geradas na validação cruzada. A Figura 5.14 ainda destaca as curvas com pior e melhor desempenho, baseado em SP, de Macaé.



(a) Modelo TCS TCC



(b) Modelo TCS TCC ELISA

Figura 5.14: Gráficos de Curva ROC para a fazenda Macaé, COM Anérgicos no Treinamento

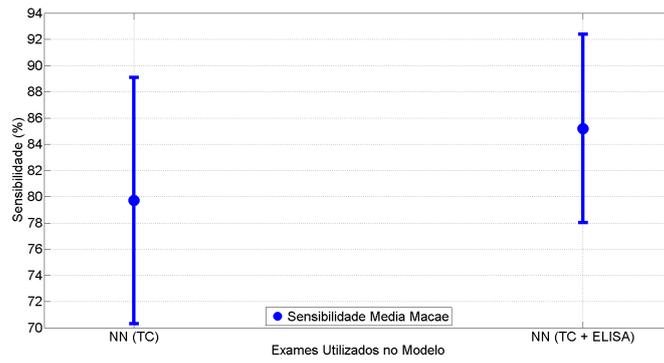
Nas curvas ROC, o primeiro critério de análise adotado é a área abaixo da curva, onde quanto maior for a mesma, melhor é o desempenho da rede representada pela curva. Considerando primeiramente apenas as curvas de pior e melhor SP, notou-se que as curvas do modelo *TCS TCC ELISA* apresentaram uma maior área do que as curvas do modelo *TCS TCC*. Para as demais curvas também nota-se esta tendência, porém em menor intensidade e que é confirmada pelos valores médios SP, onde um maior valor da medida está relacionado a uma maior área sob a ROC.

Após a análise das áreas das curvas ROC, através das mesmas são determinadas as medidas de performance deste estudo, a Sensibilidade, a Especificidade e Acuidade de cada um dos modelos. A Figura 5.15 apresenta graficamente os valores das médias e desvio padrão de cada uma das medidas de performance consideradas, somente para animais não anérgicos.

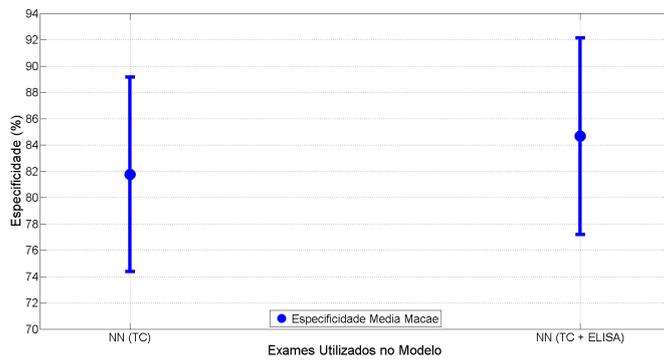
A Figura 5.15(a) destaca as Sensibilidades dos modelos neurais de Macaé. Para o modelo *TCS TCC* foi obtido uma Sensibilidade média de 89% e um desvio padrão de 9% para a medida. No modelo *TCS TCC ELISA* o seu valor médio foi 85% e um desvio padrão de 7%. A adição de ELISA neste caso acarretou em um aumento de 5 pontos percentuais (6%) e em uma diminuição de 2 pontos percentuais para o desvio padrão, em comparação com o modelo *TCS TCC*. Logo, a adição de ELISA junto aos testes cervicais aumentou a Sensibilidade dos modelos e com um maior impacto do que foi observado em UFF.

A medida de Especificidade foi relacionada na Figura 5.15(b) e apresentou valores médios de 82% para o modelo *TCS TCC* e 85% para *TCS TCC ELISA*. Já o desvio padrão retornado foi o mesmo em ambos os modelos. O aumento de seus valores médios indica que o ELISA ao ser usado em conjunto aos testes intradérmicos apresenta ganhos para esta medida de performance, enquanto o mesmo desvio padrão dos modelos indica que o ELISA não impacta na variabilidade da medida. Novamente, o acréscimo de ELISA gerou impactos positivos da Especificidade, como visto também para a fazenda UFF.

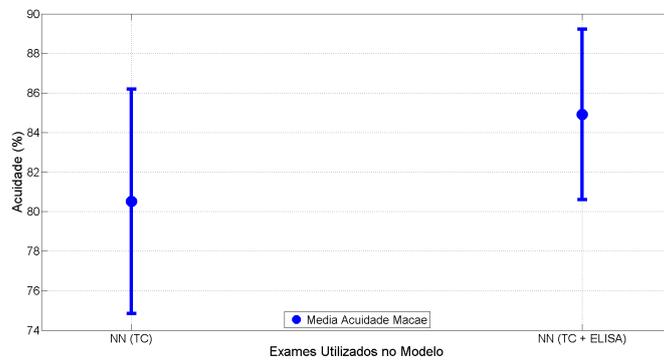
A Acuidade e a medida SP são expostas na Figuras 5.15(c) e Figuras 5.15(d), respectivamente. Ambas medidas dependem da Sensibilidade e da Especificidade para serem calculadas e conseqüentemente, acompanham os padrões observados nestas medidas. A medida SP, já analisada e também vista na Figura 5.13, tem seu aumento de valor médio decorrente ao aumento da média das duas medidas anteriores. O valor médio da Acuidade para o modelo *TCS TCC* foi de 81% e ao adicionar o ELISA (modelo *TCS TCC ELISA*) sua média aumentou para 85%. O desvio padrão da Acuidade foi de 6% para o modelo com apenas os testes cervicais e em seguida, de 4% para o modelo com ELISA. O aumento da média e a queda no desvio padrão destas medidas indicam que o modelo com ELISA apresenta resultados mais consistentes, gerando assim aumento da capacidade de classificação da TB.



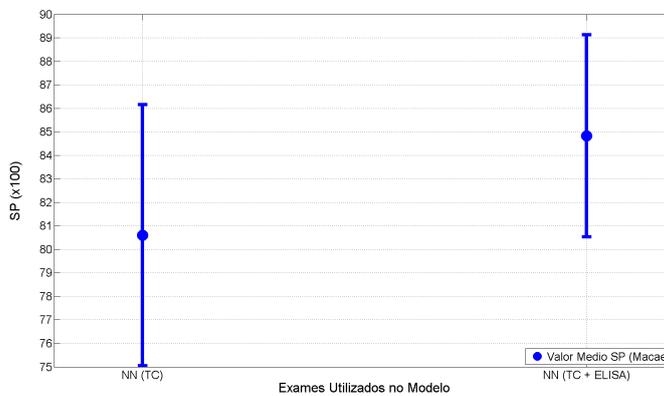
(a) Sensibilidade



(b) Especificidade



(c) Acuidade



(d) Valor SP

Figura 5.15: Média e Desvio Padrão das medidas de performance para a fazenda Macaé, COM anérgicos no treinamento

A Tabela 5.10 apresenta os resultados da performance dos modelos, apenas para os animais não anérgicos, vistos na Figura 5.15.

Tabela 5.10: Performance das redes neurais de Macaé para animais não anérgicos. Treinamento COM anérgicos

Mod. Neural	Sens. (%)	Espec. (%)	Acc. (%)	SP
TCS TCC	$80 \pm 9$	$82 \pm 7$	$81 \pm 6$	$0,81 \pm 0,06$
TCS TCC ELISA	$85 \pm 7$	$85 \pm 7$	$85 \pm 4$	$0,85 \pm 0,04$

Na análise da performance dos modelos para os animais anérgicos apenas, a Figura 5.16 apresenta graficamente a média e o desvio padrão da Sensibilidade dos modelos. O modelo *TCS TCC* apresentou uma média de 7% de Sensibilidade e um desvio padrão igual a 18%. Já o modelo *TCS TCC ELISA* possui uma Sensibilidade média de 48% com um desvio padrão de 20% da medida.

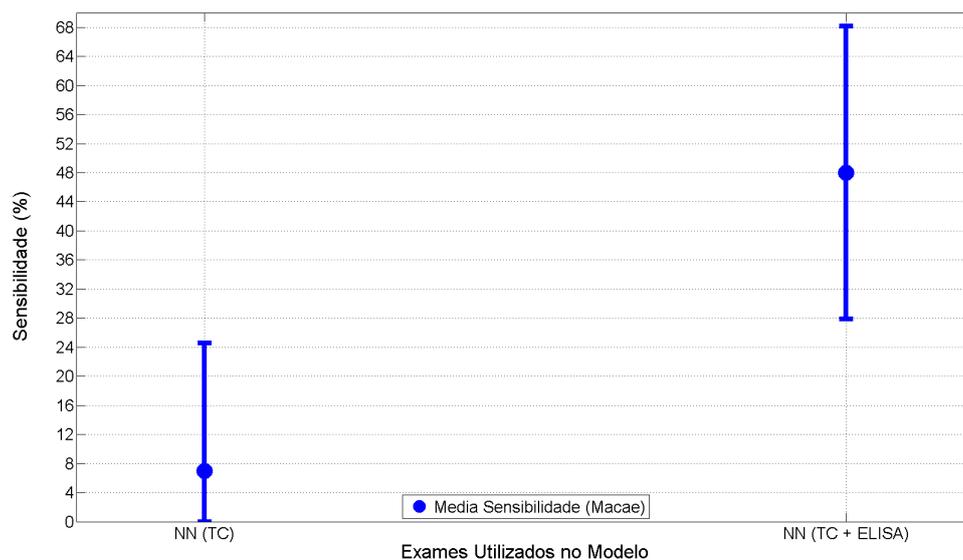


Figura 5.16: Sensibilidade dos animais anérgicos para os modelos da fazenda Macaé COM animais anérgicos no treinamento

O primeiro fator observado foi a mudança que o treinamento com animais anérgicos causou na identificação deste tipo de animal. Ao comparar as Figuras 5.12 e 5.16 o que se viu foram resultados mais consistentes quando os modelos neurais foram treinados com os animais anérgicos, uma vez que houve um aumento de seus valores médios e para o modelo *TCS TCC ELISA* houve uma diminuição de seu desvio padrão. Em ambos os treinamentos, o desvio padrão visto em todos os casos são grandes e mostram grande variação dos modelos, que é explicado pela baixa ocorrência dos eventos anérgicos já que a fazenda possui apenas 4 animais e cada

animal classificado corretamente representa um acréscimo de 25% em uma rede neural. Ao considerar neste cenário a adição do ELISA aos testes cervicais os resultados do modelo *TCS TCC ELISA* comparado ao *TCS TCC* indica um impacto positivo do exame na classificação destes animais, confirmando na fazenda o uso do ELISA na identificação de anérgicos.

A Tabela 5.11 apresenta as eficiências dos modelos neurais para os animais anérgicos vistos na Figura 5.16.

Tabela 5.11: Performance das redes neurais de Macaé para animais anérgicos. Treinamento **COM** anérgicos

Mod. Neural	Sensibilidade (%)
TCS TCC	[0 ; 25]
TCS TCC ELISA	48 ± 20

Ao considerar os animais anérgicos no treinamento dos modelos, se verificou que a adição de ELISA apresentou ganhos na identificação da Tuberculose tanto para os animais não anérgicos como para os animais anérgicos, onde os impactos observados foram maiores.

### Exames Isolados

A análise de exames isolados em Macaé busca analisar o uso de cada exame isoladamente sobre os bovinos da fazenda e verificar seu desempenhos, além de complementar as análises feitas sobre os modelos neurais desenvolvidos. A Tabela 5.12 exhibe a relação das eficiências de cada um dos exames presentes na fazenda Macaé. Foram analisados os grupos de animais não anérgicos e anérgicos separadamente. Suas performances foram calculadas a partir dos resultados de cada exame na fazenda.

Tabela 5.12: Performance dos exames isolados de Macaé para animais não anérgicos

Exame	Sens. (%)	Espec. (%)	Acc. (%)	SP
TCS	83	64	73	0,73
TCC	57	82	69	0,69
ELISA	13	77	44	0,38

Os resultados para os testes cervicais mostram que tanto o TCS quanto o TCC não são capazes de cobrir a total identificação dos animais TB+ e TB- na fa-

zenda. O TCS apresentou uma Sensibilidade e uma Especificidade de 83% e 64%, respectivamente, enquanto o TCC apresentou 57% e 82% para as mesmas medidas de performance. Os valores de Acuidade e SP acompanham a Sensibilidade e Especificidade dos testes e seus valores não indicam a identificação total da doença.

Para o ELISA, as suas medidas de performance foram de 13% para a Sensibilidade, 77% para a Especificidade, 44% de Acuidade e seu valor SP foi calculado em 0,38. Tais resultados mostram que o ELISA não é capaz de identificar a Tuberculose corretamente em todos os bovinos, mas a maior Especificidade do ELISA é confirmada com estes resultados, corroborando os resultados visto na fazenda UFF, e explicando o ganho desta medida nos modelos neurais ao adicionar o ELISA. Em Macaé, dos 22 animais TB– presentes na fazenda, 17 (77% do total de bovinos sadios) foram classificados corretamente pelo ELISA, e dentre estes, 4 foram classificados como doentes pelos dois testes cervicais. Isso mostra que o ELISA é capaz de identificar animais que os testes cervicais não conseguem e explica o ganho de Especificidade pelos modelos neurais.

A Tabela 5.13 apresenta a eficiência de cada um dos exames para os animais anérgicos.

Tabela 5.13: Performance dos exames isolados de Macaé para animais anérgicos

Mod. Neural	Sensibilidade (%)
TCS	0
TCC	0
ELISA	100

Os testes cervicais não foram capazes de identificar nenhum dos 4 animais anérgicos da fazenda. Tal fato era esperado uma vez que a anergia é caracterizada por uma fraca reação do animal infectado por *M. bovis* a estes testes. Usando como comparativo os modelos neurais desenvolvidos, tanto no cenário de treinamento dos modelos sem os animais anérgicos como no cenário de treinamento com os animais anérgicos, os modelos neurais que usavam apenas os testes cervicais não foram capazes de classificar corretamente todos os animais.

Já o ELISA, assim como na fazenda UFF, apresentou 100% de Sensibilidade na identificação destes animais. O aumento da identificação dos anérgicos também foi visto nos modelos neurais tanto para os dois cenários de treinamento dos modelos neurais. Porém, nenhum dos modelos neurais desenvolvidos, possuem capacidade de detecção de anérgicos tão eficazes quanto o exame ELISA isolado, reforçando a aplicação de ELISA para este propósito.

O experimento conduzido em Macaé mostrou que o uso de ELISA como exame complementar aos testes cervicais apresenta benefícios na identificação da doença. O ELISA consegue identificar animais anérgicos presentes na fazenda, além de apresentar boa Especificidade para os animais não anérgicos, conseguindo identificar animais TB— que os testes cervicais não identificam. A abordagem multi-diagnóstica com uso de redes neurais se mostrou vantajosa na fazenda, conseguindo identificar corretamente mais animais do que os exames aplicados separadamente.

### 5.1.3 Fazenda RS

Para a fazenda RS foram treinados apenas 2 modelos neurais. Por não possuir outros exames e por não haver qualquer informação a respeito de animais anérgicos, foi considerada apenas a modalidade de treinamento sem animais anérgicos. Os modelos considerados foram:

- *TCS TCC*
- *TCS TCC ELISA*

O primeiro fator analisado na fazenda RS foi o valor SP de seus modelos neurais. Os histogramas mostrados na Figura 5.17 trazem a distribuição destes valores SP, para cada um dos modelos neurais desenvolvidos na fazenda RS. Novamente, esta distribuição foi obtida a partir das redes neurais geradas na validação cruzada, onde cada valor está associado a uma rede neural. A figura ainda destaca a média e o desvio padrão da medida em cada modelo neural.

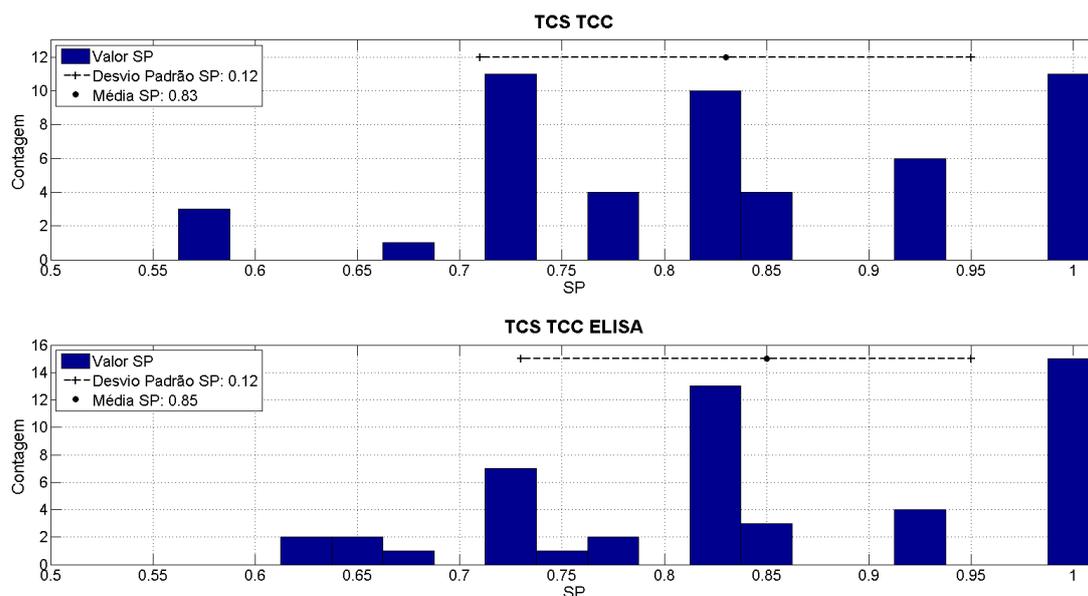
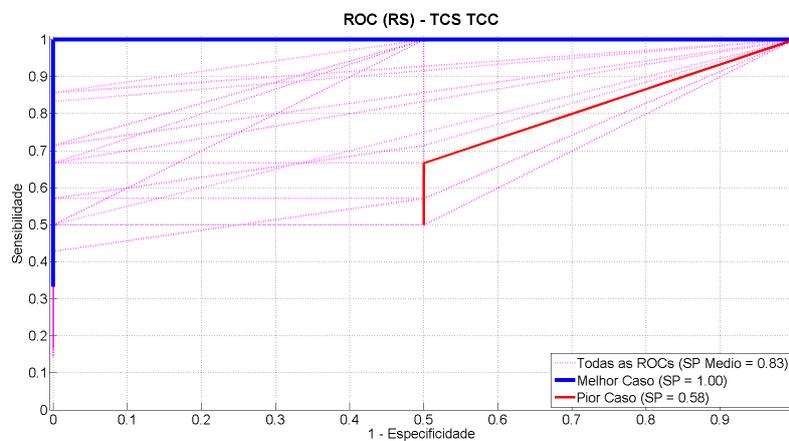


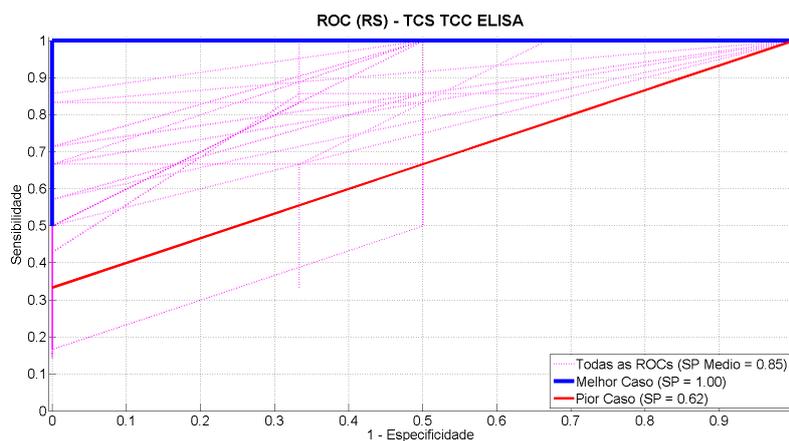
Figura 5.17: Distribuição do SP para os modelos da fazenda RS SEM animais anérgicos no treinamento

O histograma para o modelo *TCS TCC* apresentou uma média de SP igual a 0,83 e um desvio padrão de 0,12. Para o modelo *TCS TCC ELISA*, o seu histograma apresenta medidas próximas ao anterior, com média igual a 0,85 e desvio padrão de 0,12. O primeiro fator observado pelos histogramas é o alto valor de desvio padrão, para a baixa quantidade de animais da fazenda (30 bovinos), que é refletida em uma dispersão por parte dos resultados das redes. O aumento da média do SP, visto nas fazendas UFF e Macaé, é representada por um deslocamento á direita do histograma com ELISA, o que porém não é visualizado na Figura 5.17.

A Figura 5.18 apresenta as curvas ROC dos modelos e destaca para cada um destes, as curvas de pior e maior SP. Uma vez que se tem valores dispersos de SP, espera-se o mesmo tipo de comportamento das curvas ROC.



(a) Modelo TCS TCC



(b) Modelo TCS TCC ELISA

Figura 5.18: Gráficos de Curva ROC para a fazenda RS, SEM anérgicos no treinamento

O gráfico mostra de fato tal característica, tanto para o modelo *TCS TCC* quanto para o modelo *TCS TCC ELISA*. A análise pelo critério da área abaixo da curva apresenta as conclusões obtidas para a observação sobre os valores SP, onde as curvas de ambos os modelos apresentaram comportamento similar, que é refletido

nos valores SP destacados em cada curva. Em ambos os modelos as curvas de melhor SP apresentaram 1 para a medida, enquanto as curvas de pior SP apresentaram valores próximos, com 0,58 para o modelo *TCS TCC* e 0,62 para *TCS TCC ELISA*. Para as demais curvas, tal similaridade é traduzida pelo valor médio de ambas, como visto na Figura 5.17.

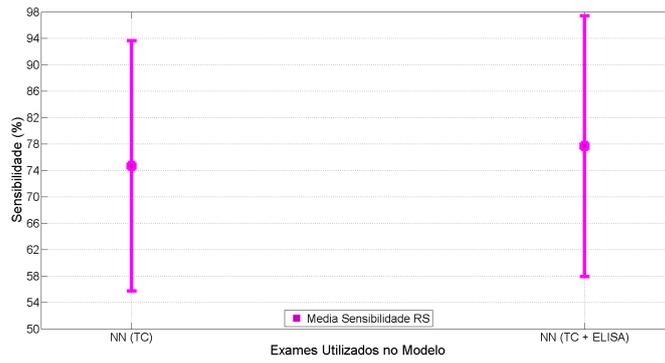
A partir das curvas ROC, foi possível calcular as medidas de Sensibilidade, Especificidade e Acuidade. Tais medidas são relacionadas graficamente na Figura 5.19 e também a partir delas, a performance dos modelos neurais foi avaliada.

A Figura 5.19(a) relaciona a Sensibilidade na fazenda RS, onde foi encontrada uma média 75% para o modelo *TCS TCC* e 78% no *TCS TCC ELISA*, além de seus desvios padrão serem, respectivamente, 19% e 20%. A adição de ELISA gera um aumento na Sensibilidade média da fazenda, porém o grande desvio padrão mostra que tal medida apresenta grande variação entre as redes neurais de cada modelo. Uma vez que suas médias são próximas e suas dispersões são altas, os dois modelos são considerados estaticamente equivalentes para Sensibilidade.

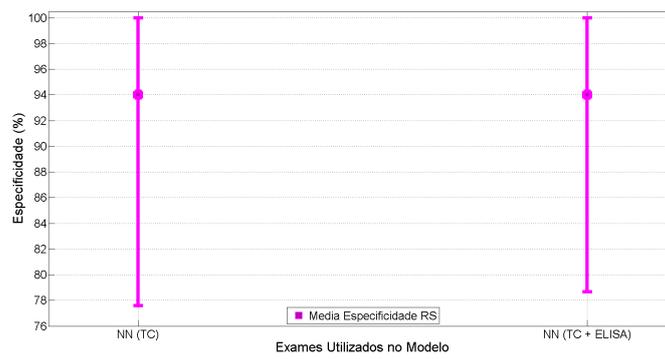
A Especificidade de RS é relacionada pela Figura 5.19(b), tendo apresentado valores médios idênticos a ambos os modelos neurais, com 94% de eficiência para esta medida. O desvio padrão de *TCS TCC* é de 16% e para o modelo *TCS TCC ELISA*, seu valor é de 15%. Os valores encontrados nos dois modelos neurais mostram que apesar da alta média da medida, há também uma grande variação por parte das redes neurais. Os modelos neurais são considerados estatisticamente empatados para a Especificidade, indicando que a adição de ELISA não apresentou impactos.

A Acuidade na fazenda RS é apresentada pela Figura 5.19(c) e para o caso específico da fazenda RS, ela apresenta um comportamento próximo da medida de Sensibilidade. A Acuidade depende tanto da Sensibilidade quanto da Especificidade para ser calculada, porém a fazenda RS apresentada quantidades desbalanceadas de animais TB+ e TB-, o que faz com que a Sensibilidade tenha maior peso para o valor final da medida. Para o modelo *TCS TCC* o valor médio encontrado foi de 78% de eficiência e um desvio padrão de 15% na mesma. No modelo *TCS TCC ELISA*, o seu valor médio foi de 80% e seu desvio padrão de 15%. Os resultados encontrados para os modelos, indicam novamente que ambos estão estatisticamente empatados, implicando que a adição de ELISA também apresenta pouco impacto para a medida de Acuidade.

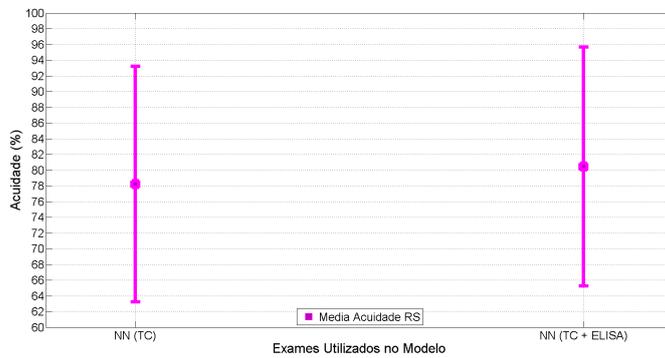
A Figura 5.19(d) relaciona a medida de SP de uma forma diferente à que foi abordada nos histogramas (Figura 5.17). Os valores de média e desvio padrão dos modelos neurais já foram abordados e foi observado que seus resultados são muito similares. A análise para esta medida é que os dois modelos neurais estão estatisticamente empatados para a medida SP, implicando que o uso de ELISA como exame complementar não apresentou impactos na fazenda RS.



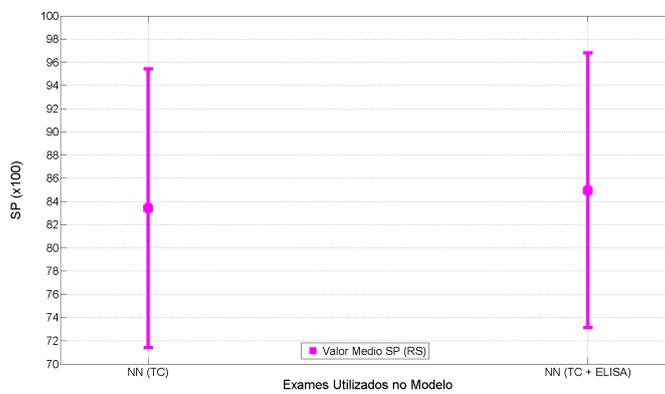
(a) Sensibilidade



(b) Especificidade



(c) Acuidade



(d) Valor SP

Figura 5.19: Média e Desvio Padrão das medidas de performance para a fazenda RS, SEM anérgicos no treinamento

A Tabela 5.14 relaciona as medidas de desempenho, vistas na Figura 5.19, para cada um dos modelos neurais da fazenda RS.

Tabela 5.14: Performance das redes neurais de RS para animais não anérgicos. Treinamento SEM anérgicos

Mod. Neural	Sens. (%)	Espec. (%)	Acc. (%)	SP
TCS TCC	75 ± 19	[78 ; 100]	78 ± 15	0,83 ± 0,12
TCS TCC ELISA	78 ± 20	[79 ; 100]	80 ± 15	0,85 ± 0,12

Pelos resultados observados pelos modelos, a adição de ELISA aos testes cervicais não apresentou um impacto na classificação dos modelos neurais desenvolvidos. Não foi observado aumento ou diminuição significativa para nenhuma das medidas de performance consideradas e o alto valor de desvio padrão das mesmas indicou que os modelos podiam ser considerados empatados estatisticamente para todas as medidas.

### Exames Isolados

A análise sobre os exames isolados em RS é um complemento às já realizadas para os modelos neurais desenvolvidos na fazenda. Nesta etapa, cada exame disponível na fazenda é avaliado separadamente segundo as mesmas medidas de performance consideradas neste estudo. A Tabela 5.15 apresenta as eficiências de cada exame da fazenda, somente para os animais não anérgicos.

Tabela 5.15: Performance dos exames isolados de RS para animais não anérgicos

Exame	Sens. (%)	Espec. (%)	Acc. (%)	SP
TCS	88	20	77	0,48
TCC	40	40	40	0,40
ELISA	92	20	80	0,49

Considerando primeiramente somente os testes cervicais, verificou-se que o TCS apresentou 88% de eficiência para a Sensibilidade, 20% para a Especificidade, 77% de Acuidade e 0,48 para a medida SP. Para o TCC, as performances encontradas foram de 40% para a Sensibilidade, a Especificidade e também para a Acuidade, enquanto sua medida SP foi de 0,40. Tais resultados mostram que os dois testes cervicais também não são capazes de identificar corretamente todos os animais da fazenda. No entanto, o comportamento destes testes apresentou similaridades com os resultados vistos nas fazendas UFF e Macaé, como por exemplo o TCS apresentando

uma maior Sensibilidade que o TCC, e este segundo teste apresentando uma maior Especificidade.

Para o ELISA, tal exame apresentou uma medida de 92% de Sensibilidade, 20% para Especificidade, 80% para sua Acuidade e um valor SP igual a 0,49. Os resultados de ELISA apresentaram diferenças, se comparado s fazendas UFF e Macaé, sendo a primeira delas o fato de sua Sensibilidade ser maior do que para os exames cervicais, fato que ocorreu nas demais fazendas. Outro ponto foi a sua baixa Especificidade, diferente do que fora observado anteriormente, onde o exame apresentou medidas de Especificidade maiores do que sua Sensibilidade. Uma razão para esta performance de ELISA foi a escolha de seu ponto de corte baseado em um grupo de controle de animais TB-, privilegiando a Sensibilidade do exame.

Assim, o experimento conduzido na fazenda RS mostrou que a adição do ELISA como exame complementar não gerou impactos no correto diagnóstico da doença, onde o seu uso em conjunto aos teste cervicais apresentou as mesmas performances vistas apenas para estes exames.

#### **5.1.4 Análise conjunta das fazendas**

Nesta etapa do experimento, busca-se analisar o impacto da implantação de novos exames ao diagnóstico da tuberculose, porém com uma visão mais abrangente sobre todas as fazendas. Os resultados obtidos para cada uma das fazendas nas seções 5.1.1, 5.1.2 e 5.1.3 são comparados e analisados juntos.

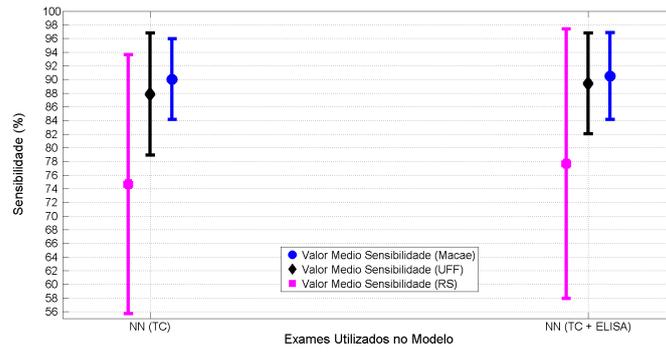
Considerou-se nesta análise apenas os modelos neurais desenvolvidos com os exames presentes nas três fazendas, sendo assim, os modelos usados são:

- *TCS TCC*
- *TCS TCC ELISA*.

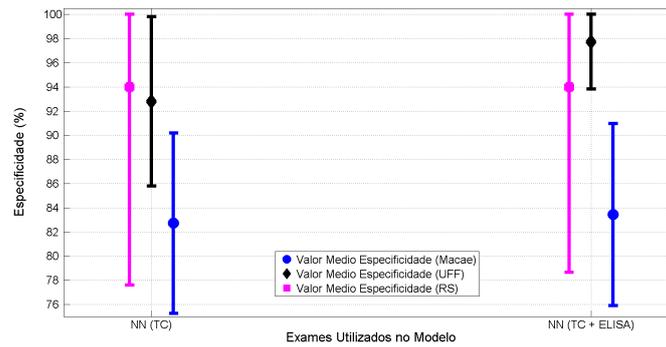
Apesar da ausência de animais anérgicos na fazenda RS, usou-se esta fazenda na comparação dos modelos com Macaé e UFF nos cenários em que elas foram treinadas com e sem animais anérgicos. Na análise feita somente para os animais anérgicos, considerou-se apenas as fazendas UFF e Macaé.

#### **Modelos treinados sem animais anérgicos**

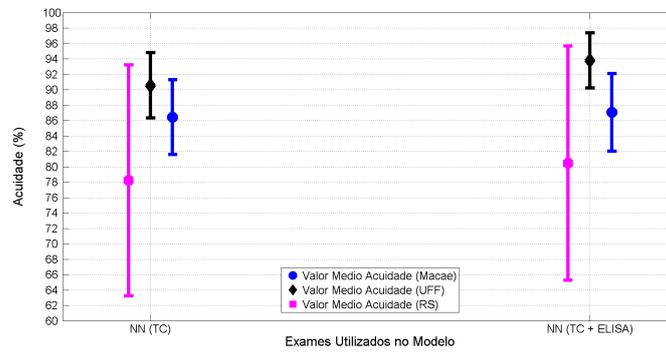
A primeira análise é feita para o cenário em que não havia animais anérgicos no treinamento dos modelos neurais. A Figura 5.20 contém a média e o desvio padrão das medidas de performance consideradas, para cada um dos modelos neurais nas fazendas UFF, Macaé e RS. A figura relaciona apenas os animais não anérgicos de cada fazenda.



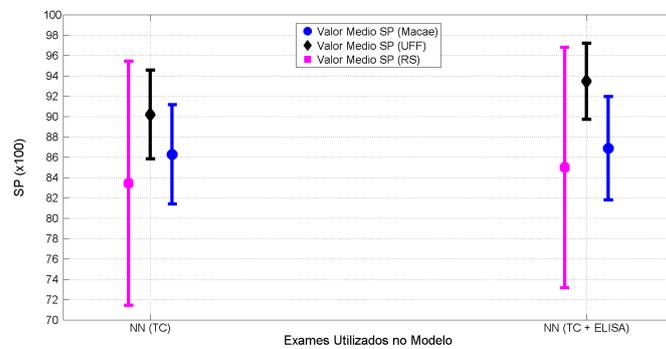
(a) Sensibilidade



(b) Especificidade



(c) Acuidade



(d) Valor SP

Figura 5.20: Média e Desvio Padrão das medidas de performance para todas as fazendas, SEM anérgicos no treinamento

A Figura 5.20(a) compara a medida de Sensibilidade dos modelos das três fazendas. A figura mostra que as fazendas UFF e Macaé possuem menor desvio padrão quando comparadas com RS, o que indica que os modelos neurais de ambas apresentam uma menor variação estatística para a medida e conseqüentemente, resultados mais confiáveis. A fazenda Macaé possui os maiores valores médios de Sensibilidade, com UFF apresentando performance muito similar e por fim, a fazenda RS com seus valores médios distantes das demais fazendas. Logo, as fazendas UFF e Macaé podem ser consideradas estatisticamente equivalentes em ambos os modelos. No entanto se observa que em todas as fazendas a adição de ELISA não causa impacto na Sensibilidade dos modelos.

A medida de Especificidade é relacionada pela Figura 5.20(b). Novamente ao comparar as três fazendas, nota-se que UFF e Macaé possuem um menor desvio padrão se comparadas à fazenda RS, e conseqüentemente estas duas fazendas possuem uma menor variação de seus modelos neurais. As fazendas UFF e RS possuem valores médios de Especificidade muito próximos mas no caso de RS, seus modelos apresentaram uma alta dispersão, o que indica que sua média foi gerada a partir de provável erro estatístico do modelo. A fazenda UFF foi a única a apresentar aumentos na Especificidade com a adição de ELISA, enquanto os modelos *TCS TCC* e *TCS TCC ELISA*, dentro das próprias fazendas de Macaé e RS, são estatisticamente equivalentes. Logo, foi observado que apenas na fazenda UFF a adição de ELISA apresentou ganhos na classificação de Tuberculose.

A Acuidade é relacionada na Figura 5.20(c) e nela foi observado que a fazenda UFF apresentou os melhores resultados para esta medida, com as maiores médias e os menores desvios padrão em seus modelos. Em contrapartida, a fazenda RS apresentou os valores médios mais baixos e os maiores valores de desvio padrão em seus modelos. A fazenda UFF foi a única que apresentou aumento em seu valor médio quando o ELISA foi adicionado ao modelo neural, enquanto as fazendas de Macaé e RS não apresentaram mudanças significativas e são consideradas como estatisticamente equivalentes. Logo, apenas em UFF a adição de ELISA acarretou em ganho de performance na Acuidade.

A Figura 5.20(d) apresenta os valores da medida SP nas três fazendas. O valor SP apresenta um comportamento similar ao visto na Acuidade, com os valores médios de UFF sendo os maiores e com a fazenda RS com os menores valores médios, apesar de serem valores mais próximos de Macaé do que foi observado na Acuidade. Sendo o SP a principal medida de avaliação de performance dos modelos neurais, apenas os modelos de UFF apresentaram aumento na eficiência de classificação ao se adicionar o ELISA. Para Macaé e RS, a adição de ELISA não apresenta impactos e os modelos de cada fazenda são considerados equivalentes.

A Tabela 5.16 agrupa os valores das medidas de performance vistas na Figura

5.20 e esta tabela é composta pelos valores já observados nas Tabelas 5.1, 5.8 e 5.14.

Tabela 5.16: Comparação dos modelos neurais entre todas as fazendas, somente para animais não anérgicos. Treinamento dos modelos SEM anérgicos

Mod. Neural	Faz.	Sens. (%)	Espec. (%)	Acc. (%)	SP
TCS TCC	UFF	88 ± 9	93 ± 7	91 ± 4	0,90 ± 0,04
TCS TCC ELISA	UFF	89 ± 7	[94 ; 100]	94 ± 4	0,93 ± 0,04
TCS TCC	Macaé	90 ± 6	83 ± 7	86 ± 5	0,86 ± 0,05
TCS TCC ELISA	Macaé	91 ± 6	83 ± 8	87 ± 5	0,87 ± 0,05
TCS TCC	RS	75 ± 19	[78 ; 100]	78 ± 15	0,83 ± 0,12
TCS TCC ELISA	RS	78 ± 20	[79 ; 100]	80 ± 15	0,85 ± 0,12

Para a análise dos animais anérgicos, apenas as fazendas UFF e Macaé têm seus valores comparados. A medida considerada nesta análise é apenas a Sensibilidade. A Figura 5.21 relaciona graficamente as Sensibilidades dos modelos *TCS TCC* e *TCS TCC ELISA* em ambas as fazendas.

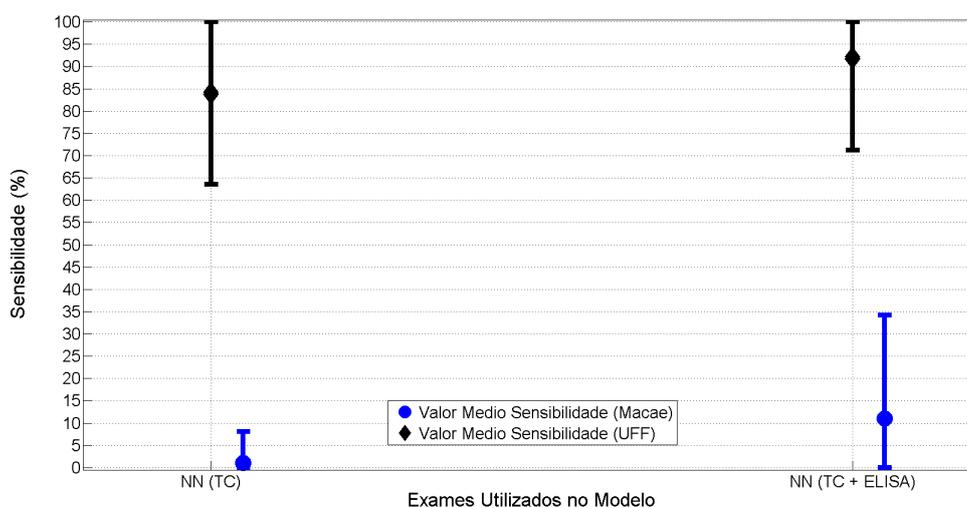


Figura 5.21: Distribuição do SP para os modelos da fazenda RS SEM animais anérgicos no treinamento

A Sensibilidade média de ambas fazendas são totalmente distintas, com a fazenda UFF apresentando boas Sensibilidades médias de 84% e 92% para os modelos *TCS TCC* e *TCS TCC ELISA*, respectivamente, enquanto Macaé apresentou valores de 1% e 11% para os seus modelos. O valores de desvio padrão em UFF são de 20% e 21%, enquanto Macaé apresentou desvios de 7% e 23%, respectivamente para os modelos *TCS TCC* e *TCS TCC ELISA*. A fazenda UFF consegue identificar os animais anérgicos com maior eficiência do que Macaé, ainda que Macaé possua 1 animal a mais que UFF, o que influencia diretamente na taxa de Sensibilidade

dos modelos. Porém em ambas as fazendas, a adição do ELISA gerou aumentos na Sensibilidade média dos anérgicos, tal como aumento nos seus desvios padrão, que podem ser explicados pela baixa ocorrência desses animais nas fazendas. Logo, em ambas as fazendas a adição de ELISA gerou impacto positivo na classificação de anérgicos. Tais resultados mostram que a aplicação específica de ELISA aos anérgicos não está restrita a apenas uma fazenda.

A Tabela 5.17 relaciona os valores de Sensibilidade vistos na Figura 5.21, para cada fazenda e modelo neural. Seus valores são compostos pelos valores observados nas Tabelas 5.2 e 5.9.

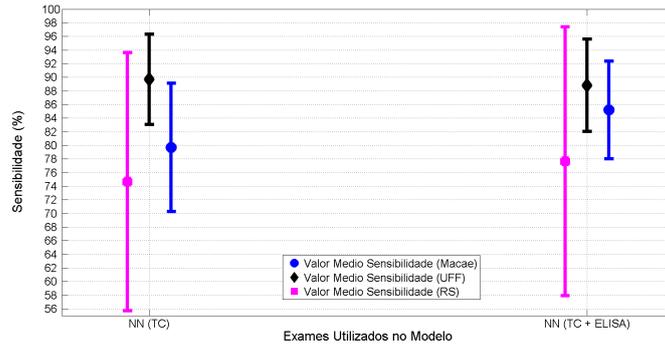
Tabela 5.17: Comparação dos modelos entre as fazendas UFF e Macaé, somente animais anérgicos. Treinamento dos modelos SEM anérgicos

Mod. Neural	Fazenda	Sensibilidade (%)
TCS TCC	UFF	[64 ; 100]
TCS TCC ELISA	UFF	[71 ; 100]
TCS TCC	Macaé	[1 ; 8]
TCS TCC ELISA	Macaé	[0 ; 34]

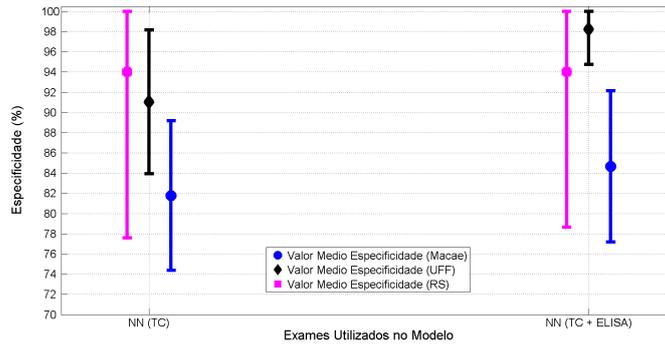
Portanto, na comparação das três fazendas, ao considerar o treinamento sem animais anérgicos, verificou-se que a adição de ELISA foi benéfica apenas para as fazendas de Macaé e UFF, onde a primeira apresentou ganhos na identificação somente para os animais anérgicos, enquanto a segunda apresentou ganhos também para na classificação de animais não anérgicos.

### Modelos treinados com animais anérgicos

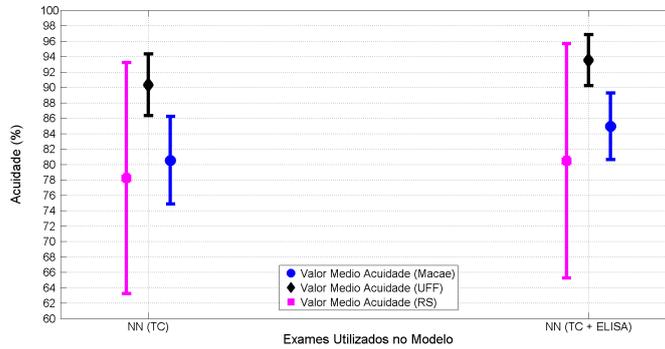
Para o segundo cenário proposto foi feita uma análise sobre as três fazendas juntas, porém com as fazendas UFF e Macaé sendo treinadas com seus animais anérgicos. A Figura 5.22 apresenta a média e o desvio padrão para as medidas de performance de Sensibilidade, Especificidade, Acuidade e valor SP apenas para os seus animais não anérgicos. Os valores apresentados na figura são as medidas de performance já vistas em cada fazenda e foram retirados das Tabelas 5.3, 5.10 e 5.14.



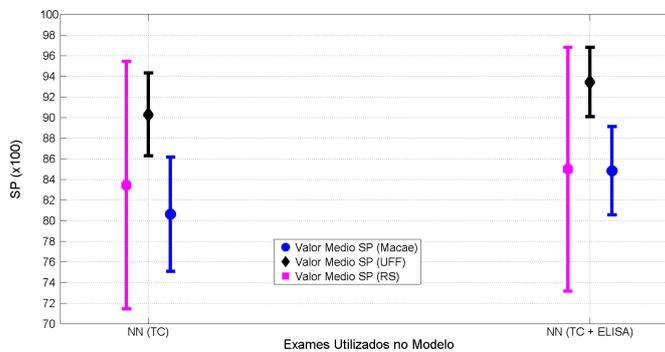
(a) Sensibilidade



(b) Especificidade



(c) Acuidade



(d) Valor SP

Figura 5.22: Média e Desvio Padrão das medidas de performance para todas as fazendas, COM anérgicos no treinamento

A Sensibilidade dos modelos é apresentada na Figura 5.22(a). A fazenda UFF apresenta pouca mudança ao adicionar ELISA ao modelo, uma vez que seu desvio padrão não se altera e que há uma queda de 1 ponto percentual em seu valor médio, o que no entanto não é relevante e indica que os modelos são estatisticamente equivalentes. A fazenda RS apresenta um aumento na Sensibilidade média com o uso de ELISA, aumentando de 75% para 78%, o que no entanto não é significativo dado o seu alto desvio padrão, indicando que o efeito do novo exame não é relevante ao modelo. A fazenda Macaé apresenta ganhos com o ELISA, com sua média aumentando e seu desvio padrão diminuindo, o que indica maior consistência do classificador. Apesar de UFF apresentar em ambos os modelos uma melhor performance, apenas a fazenda Macaé apresentou ganhos na Sensibilidade com a inclusão de ELISA.

A Especificidade dos modelos das três fazendas é relacionada pela Figura 5.22(b). A fazenda RS apresenta no modelo *TCS TCC* o maior valor médio de Especificidade, no entanto, também possui uma grande variação na medida, com seu desvio padrão sendo igual a 19% e com a adição de ELISA não há impacto significativo, o que indica que os modelos de RS são estatisticamente equivalentes. Para UFF e Macaé, a adição de ELISA apresenta impactos positivos, com o aumento dos valores médios de ambas, com a diminuição do desvio padrão em UFF e por fim, com o desvio padrão de Macaé mantendo seu valor. O uso de ELISA na fazenda UFF também apresenta um impacto maior do que na fazenda Macaé e isso pode ser explicado pelo fato de UFF possuir uma maior variedade de ELISAs (*MPB70*, *MPB83* e *ESAT-6*) além da fazenda possuir dados de testes cervicais com medições mais precisas que Macaé, o que afeta diretamente os diagnósticos finais por esses testes. Apesar das diferenças, o ELISA gera benefícios na identificação de TB apenas em UFF e Macaé.

A Acuidade e o valor SP dos modelos são expostos nas Figuras 5.22(c) e 5.22(d), respectivamente. O comportamento observado tanto para a Acuidade quanto para o SP foram os mesmos dentro de cada fazenda. A adição de ELISA gerou aumentos nos valores médios das medidas em todas as fazendas, enquanto que para os valores de desvio padrão houve queda nos valores para os dados de UFF e Macaé, mas manteve-se o desvio de 15% para a fazenda RS. Considerando a fazenda RS, os ganhos obtidos não são representativos e seus altos valores de desvio padrão mostram que, por sua grande variação nas médias, seus modelos podem ser considerados estatisticamente empatados. Ao considerar UFF e Macaé, a adição de ELISA proporcionou impactos positivos para estas medidas, indicando que os resultados dos modelos com o exame são mais consistentes e com aumento em suas eficiências.

A Tabela 5.18 reúne todas as métricas de performance dos 6 modelos neurais, analisadas na Figura 5.22.

Tabela 5.18: Comparação dos modelos neurais entre todas as fazendas, somente para animais não anérgicos. Treinamento dos modelos **COM** anérgicos

Mod. Neural	Faz.	Sens. (%)	Espec. (%)	Acc. (%)	SP
TCS TCC	UFF	90 ± 7	91 ± 7	90 ± 4	0,90 ± 0,04
TCS TCC ELISA	UFF	89 ± 7	[94 ; 100]	94 ± 3	0,93 ± 0,03
TCS TCC	Macaé	80 ± 9	82 ± 7	81 ± 6	0,81 ± 0,06
TCS TCC ELISA	Macaé	85 ± 7	85 ± 7	85 ± 4	0,85 ± 0,04
TCS TCC	RS	75 ± 19	[78 ; 100]	78 ± 15	0,83 ± 0,12
TCS TCC ELISA	RS	78 ± 20	[79 ; 100]	80 ± 15	0,85 ± 0,12

Para os animais anérgicos de UFF e Macaé, a Figura 5.23 compara o comportamento da Sensibilidade dos modelos neurais. Diferentemente do que foi visto no treinamento dos modelos sem anérgicos, o acréscimo de ELISA em Macaé apresentou um grande aumento de Sensibilidade na fazenda, fazendo o valor médio saltar 41 pontos percentuais e aumentando o desvio padrão em apenas 2%. A fazenda UFF também apresentou impactos com ELISA neste cenário, aumentando seu valor médio em 4 pontos percentuais e diminuindo o desvio padrão para 11%.

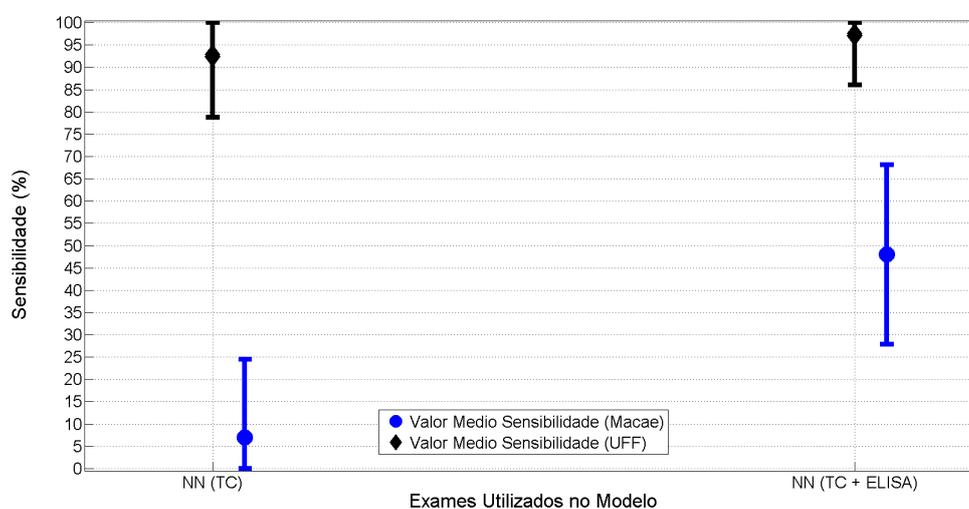


Figura 5.23: Distribuição do SP para os modelos da fazenda RS **SEM** animais anérgicos no treinamento

A Tabela 5.19 apresenta as performances apresentadas na Figura 5.23, relacionando o modelo neural, a fazenda a qual pertence e seus valores de Sensibilidade.

Tabela 5.19: Comparação dos modelos entre as fazendas UFF e Macaé, somente animais anérgicos. Treinamento dos modelos COM anérgicos

Mod. Neural	Fazenda	Sensibilidade (%)
TCS TCC	UFF	[79 ; 100]
TCS TCC ELISA	UFF	[86 ; 100]
TCS TCC	Macaé	[0 ; 25]
TCS TCC ELISA	Macaé	48 ± 20

O grande aumento de performance para a fazenda Macaé não foi observado em UFF e um fato que pode ter influenciado este resultado foi os resultados de seus testes cervicais. Em UFF, o TCC com diagnóstico inconclusivo foi encontrado apenas em animais TB+ da fazenda, com quatro ocorrências. Dos casos relatados, dois bovinos são não anérgicos e classificados como positivos pelo TCS, e os outros dois animais anérgicos, porém classificados como inconclusivos pelo TCS. O modelo neural *TCS TCC* de UFF ao ser treinado com estes dados, classificou estes animais inconclusivos com 99% de Sensibilidade média e um desvio padrão de 7%, indicando que apenas com o diagnóstico de TCC definido como inconclusivo, era possível identificar o animal corretamente como TB+. Tal fato indica que a fazenda UFF classificou os animais anérgicos não somente com ELISA, mas também com seus testes cervicais. Já na fazenda Macaé, seus resultados indicaram (Figura 5.23) que a fazenda baseou a identificação dos anérgicos com grande importância somente no ELISA. A análise feita sobre os testes cervicais apenas é mostrada na Seção 5.3.

Por fim, observa-se que apenas nas fazendas UFF e Macaé, a adição de ELISA foi traduzido em poder de classificação para as mesmas. Considerando os animais não anérgicos, a fazenda UFF obteve ganhos na Especificidade apenas, enquanto a fazenda Macaé apresentou aumentos na Especificidade e também para na Sensibilidade dos modelos. Para os animais anérgicos, novamente o ELISA se mostrou como um exame propício na identificação destes animais e em ambas fazendas o seu uso apresentou ganhos na Sensibilidade destes bovinos.

## 5.2 Repetibilidade

Para este segundo experimento a proposta foi aplicar a Repetibilidade ao experimento anterior, apresentado na Seção 5.1, e verificar seus resultados e o impacto que a adição de novos exames tem nas fazendas analisadas em condições diferentes ao primeiro estudo.

Nesse experimento, decidiu-se por propagar os dados dos animais da fazenda RS

nos modelos neurais desenvolvidos para as fazendas de UFF e Macaé. Os modelos testados são o *TCS TCC* e o *TCS TCC ELISA*, por serem os modelos que contemplam os exames disponíveis nas três fazendas. Esta seção apresenta os resultados separados em duas categorias, com e sem a presença de anérgicos nos modelos treinados, mas como RS não possui animais desse tipo, somente os não anérgicos foram usados como entrada.

### 5.2.1 Redes de Operação SEM anérgicos no treinamento

O primeiro passo para a condução deste experimento foi escolher as redes neurais de operação de cada modelo neural em cada fazenda. O parâmetro considerado na escolha destas redes foi o valor SP, onde todas elas foram avaliadas com todos os dados de suas fazendas. Essas informações foram propagadas nas redes, obtendo as performances de cada uma e em seguida foram escolhidas as redes que apresentassem a maior medida SP, em cada modelo. A Tabela 5.20 relaciona as redes de operação escolhidas neste experimento, separando-as por modelo neural e fazenda. A tabela contém as medidas de performance de Sensibilidade, Especificidade, Acuidade e SP de cada rede de operação.

Tabela 5.20: Redes de Operação para os grupos de treinamento e teste, SEM anérgicos no treinamento

Fazenda	Mod. Neural	Sens. (%)	Espec. (%)	Acc. (%)	SP
UFF	TCS TCC	90	87	88	0,88
UFF	TCS TCC ELISA	85	100	93	0,92
Macaé	TCS TCC	87	82	84	0,84
Macaé	TCS TCC ELISA	91	82	87	0,87

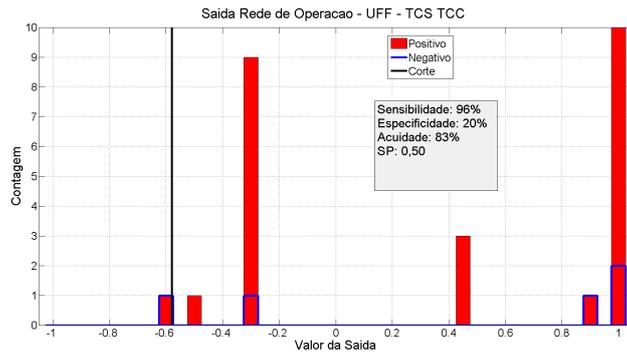
Para cada um dos modelos neurais de cada uma das fazendas, foram geradas diferentes redes neurais na etapa de validação de cruzada de seus treinamentos. Para escolher uma RNA como uma rede de operação, idealmente deve-se escolher uma rede, dentre todas geradas, baseado apenas em dados antes nunca vistos pela rede testada. No caso, deveriam ser usados conjuntos de dados de teste e validação que fossem representativos aos possíveis casos que representam o problema. No entanto, os dados disponíveis para cada fazenda são escassos e a divisão destes em grupos de validação e teste faz com que tais grupos não sejam representativos e abrangentes. Sendo assim, para contornar tal limitação foram usados todos os dados de cada fazenda, ou seja, foram incluídos os dados usados em seus treinamentos.

O passo seguinte consistiu na propagação dos dados dos animais de RS nas

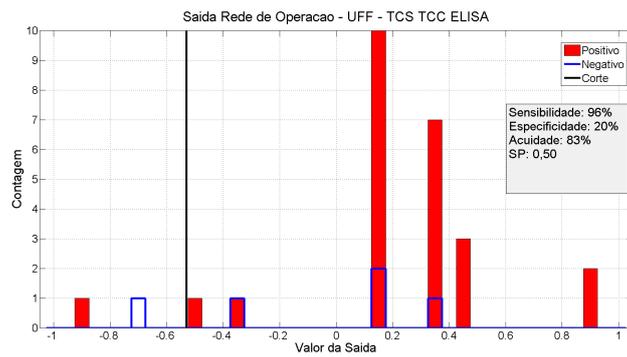
redes de operação escolhidas, ou seja, cada rede de operação classificou os animais de RS e depois foi avaliada a performance destas classificações. As redes neurais desenvolvidas neste estudo possuem um valor de ponto de corte que divide as saídas geradas em diagnósticos TB+ ou TB-, e neste caso, as redes de operação usaram os pontos de corte definidos apenas por seus conjuntos de teste. A Figura 5.24 relaciona a saída de cada rede de operação deste cenário para os dados de RS propagados, destacando ainda as medidas de Sensibilidade, Especificidade, Acuidade e SP da propagação.

As Figuras 5.24(a) e 5.24(b) apresentam as saídas de RS, respectivamente, nas redes *TCS TCC* e *TCS TCC ELISA* da fazenda UFF, e em ambas as propagações os resultados obtidos foram os mesmos para os dois modelos neurais. Considerando primeiramente a Sensibilidade, as redes de operação apresentaram uma eficiência de 96%, o que para a fazenda RS significa que apenas um animal TB+ não foi classificado corretamente. A performance de Especificidade foi de apenas 20% nas redes de UFF, onde dos cinco animais indicados como TB-, apenas um deles conseguiu ser classificado corretamente. Para a Acuidade foram obtidas performances de 83%, onde este valor foi majoritariamente conseguido pelo alto acerto de bovinos TB+. Já a medida SP foi igual a 0,50 nos modelos, determinada pela baixa Especificidade encontrada. Observando apenas as saídas e seus resultados, viu-se que o uso de ELISA não impactou na classificação de RS pelas redes de operação de UFF.

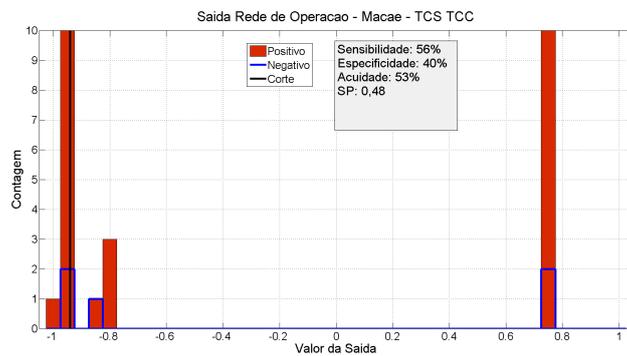
As Figuras 5.24(c) e 5.24(d) relacionam respectivamente, as saídas das rede de operação de *TCS TCC* e *TCS TCC ELISA* de Macaé. Analisando primeiramente suas Especificidades, o que se viu foi que ambas conseguiram a mesma performance, com 40% de classificação correta, representando um total de dois animais em RS. Para suas Sensibilidades os modelos apresentaram diferentes resultados, com a rede neural *TCS TCC* apresentando 56% de eficiência contra 64% da rede *TCS TCC ELISA*. O aumento de performance visto com o uso de ELISA representou um total de dois bovinos TB+ a mais classificados corretamente, também visto no experimento anterior (Seção 5.1) em Macaé, onde o uso de ELISA apresentou ganhos na Sensibilidade dos modelos. Uma vez que houve aumento de Sensibilidade e estagnação na Especificidade, as medidas de Acuidade e valor SP apresentaram aumentos com o uso de ELISA. Logo, ao propagar os bovinos nas redes de operação de Macaé, o que se viu foi um pequeno aumento na Sensibilidade das redes operação.



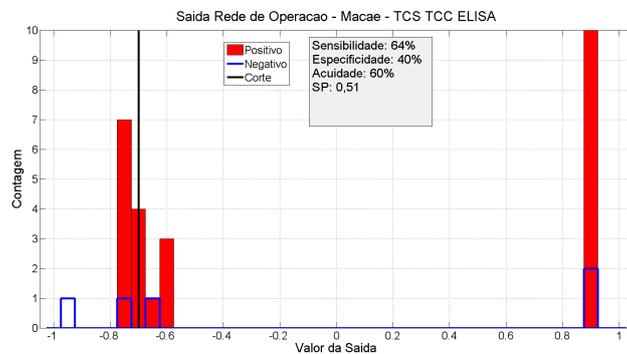
(a) Modelo TCS TCC de **UFF**



(b) Modelo TCS TCC ELISA de **UFF**



(c) Modelo TCS TCC de **Macaé**



(d) Modelo TCS TCC ELISA de **Macaé**

Figura 5.24: Saídas das Redes de Operação para os animais de RS, SEM anérgicos no treinamento

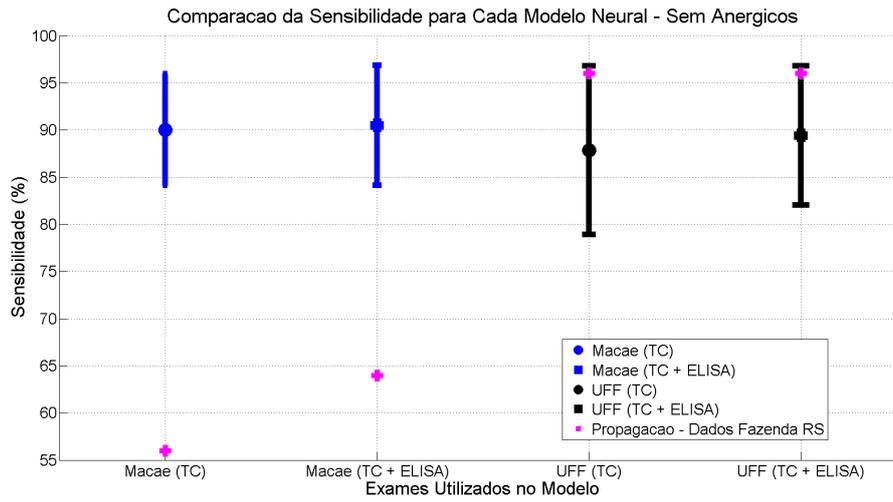
A Tabela 5.21 relaciona as medidas de performance obtidas pelas propagações dos dados de RS nas redes de operação de UFF e Macaé, vistas na Figura 5.24.

Tabela 5.21: Dados de RS propagados nas redes de operação treinadas SEM Anérgicos

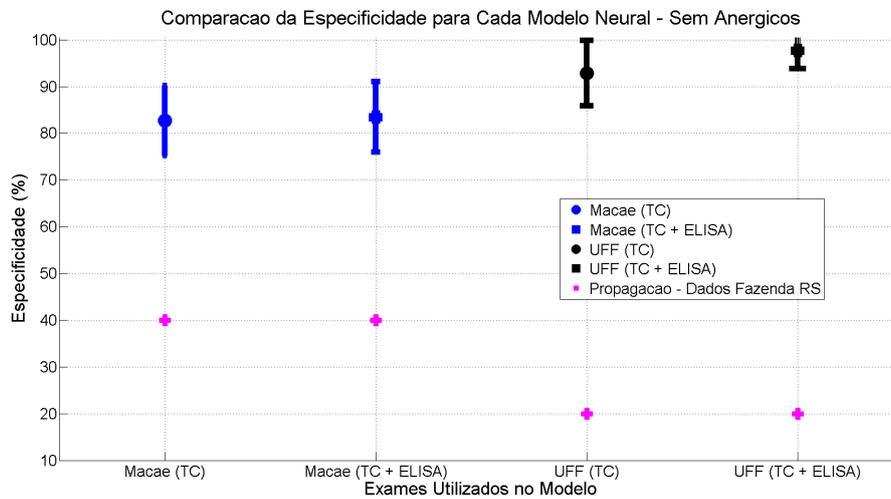
Fazenda	Mod. Neural	Sens. (%)	Espec. (%)	Acc. (%)	SP
UFF	TCS TCC	96	20	83	0,50
UFF	TCS TCC ELISA	96	20	83	0,50
Macaé	TCS TCC	56	40	53	0,48
Macaé	TCS TCC ELISA	64	40	60	0,51

Os valores SP vistos para todos os modelos indicam que a classificação dos dados de RS não apresentou boa performance e a razão para este fato ter ocorrido em cada fazenda foi diferente. A medida SP é uma medida que busca equilibrar a Sensibilidade com a Especificidade e caso uma das medidas apresentem valores muito discrepantes, ela apresenta uma queda em seu valor. No caso de UFF, a diferença vista foi 76 pontos percentuais entre as medidas, explicando assim seu baixo valor de SP. Para a fazenda Macaé, os valores encontrados foram mais baixos, porém próximos o que acabou gerando os baixos valores de SP na fazenda.

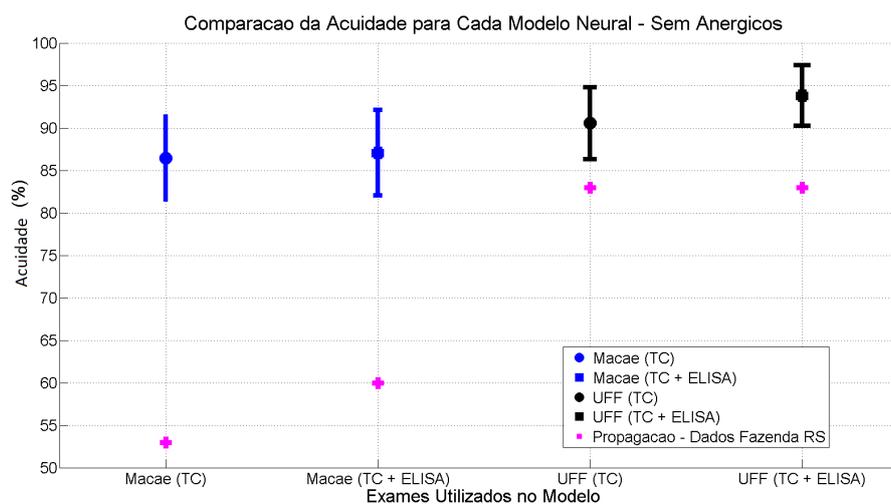
Outra forma para avaliar o desempenho da classificação de RS pelas redes de operação de UFF e Macaé é verificar se as medidas de performance se encontram dentro do erro estatístico dos modelos neurais considerados. Tal fato indica que o comportamento sobre aqueles dados está dentro do esperado para aquele modelo e, para o problema específico da Tuberculose, indica que as características e impactos já observados pela adição de ELISA aos testes cervicais se mantém. A Figura 5.25 relaciona as performances dos dados de RS propagados nas redes de operação com as medidas de performance dos modelos neurais, relacionados nas Tabelas 5.1 e 5.8. Para esta análise a medida SP foi desconsiderada, uma vez que ela atribui às medidas de Sensibilidade e Especificidade a mesma importância e para o caso de RS, as poucas ocorrências de animais TB– geram grandes impactos no SP, porém que não refletem bem o desempenho da rede. Sendo assim, a figura apenas considera as medidas de Sensibilidade, Especificidade e Acuidade.



(a) Sensibilidade



(b) Especificidade



(c) Acuidade

Figura 5.25: Dados da fazenda RS simulados nos modelos neurais de UFF e Macaé treinados SEM anérgicos

Ao comparar as saídas obtidas por cada rede operação com a performance dos modelos com sua variação estatística, viu-se que as medidas de Acuidade e Especificidade estavam fora das regiões de média e desvio padrão de todos os modelos de UFF e Macaé. A Acuidade em UFF apresentou-se próxima do comportamento dos modelos da fazenda, mas sua performance foi comprometida pela baixa Especificidade, enquanto em Macaé mesmo os animais TB+ não apresentaram uma classificação próxima do comportamento esperado da fazenda, o que também fez sua Acuidade apresentar performances abaixo dos modelos. Apenas a medida de Sensibilidade de UFF apresentou seus valores dentro do desvio padrão do modelo, indicando que para esta medida os resultados obtidos para RS estão dentro do comportamento observado nos mesmos. Tal resultado indica que, ao menos para os animais TB+ de RS, os efeitos da adição de ELISA podem ser os mesmos vistos em UFF.

### 5.2.2 Redes de Operação COM anérgicos no treinamento

Esta segunda análise considera apenas os modelos neurais de UFF e Macaé que foram treinados com seus animais anérgicos. Novamente a escolha das redes neurais de operação foi feita considerando a medida SP. Todas as redes neurais geradas na etapa de validação cruzada tiveram seus conjuntos de dados completos propagados e a Tabela 5.22 relaciona as medidas de performance das redes de operação selecionadas para estes dados.

Tabela 5.22: Redes de Operação para os grupos de treinamento e teste, COM anérgicos no treinamento

Fazenda	Mod. Neural	Sens. (%)	Espec. (%)	Acc. (%)	SP
UFF	TCS TCC	91	87	89	0,89
UFF	TCS TCC ELISA	87	100	93	0,93
Macaé	TCS TCC	74	82	78	0,78
Macaé	TCS TCC ELISA	85	82	84	0,83

Novamente, assim como no caso dos modelos treinados sem animais anérgicos (Seção 5.2.1), as redes de operação de cada um dos modelos considerados, foram escolhidas com o uso de todos os dados disponíveis nas fazendas. Neste segundo caso considerado, mesmo o acréscimo dos animais anérgicos em seus conjuntos de dados, não representou um aumento considerável na quantidade de bovinos e que pudesse fazer com que os conjuntos de validação e testes fossem representativos ao cenário do problema da Tuberculose bovina. Assim, a fim de se escolher as redes de operação baseado na maior diversidade de dados possível, usou-se também os dados

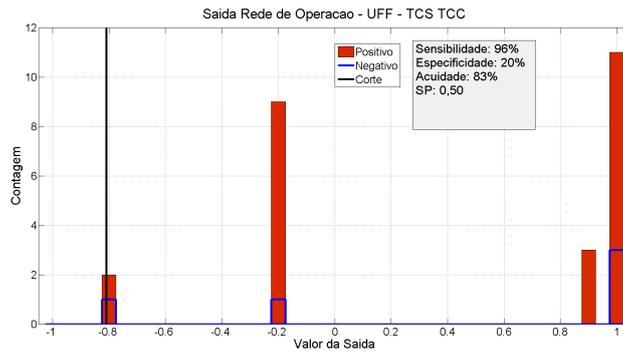
de treinamento para esta escolha.

Com a rede de operação selecionada em cada modelo neural, foi feita em seguida a propagação dos dados de RS nestas redes. O procedimento adotado foi o mesmo para o caso de treinamento sem anérgicos, com o ponto de corte da cada rede sendo definido apenas com base da performance do conjunto de teste da rede de operação. A Figura 5.26 relaciona a saída apresentada para a propagação dos dados de RS, tal como o seu ponto de corte e as medidas de performance de Sensibilidade, Especificidade, Acuidade e SP.

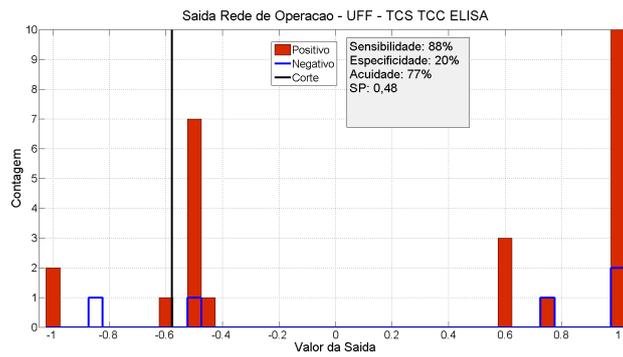
As Figuras 5.26(a) e 5.26(b) apresentam as saídas de RS nos modelos de UFF, respectivamente *TCS TCC* e *TCS TCC ELISA*. Em ambas as propagações, dos cinco animais ditos como TB–, apenas um animal (20%) foi classificado corretamente. A medida de Sensibilidade na fazenda apresentou perdas em sua eficiência ao adicionar o exame ELISA, o que significa que a rede de operação possui maiores dificuldades para classificar os animais TB+ de RS. Com esta diminuição, as medidas de Acuidade e SP também apresentaram quedas, com a primeira passando de 83% para 77% em eficiência, e o SP diminuindo de 0,50 para 0,48. O modelo *TCS TCC* classificou bovinos com saídas próximas ao valor 1 em mais ocasiões do que o modelo *TCS TCC ELISA*, e isso indica que o primeiro modelo possui maior certeza na classificação desses bovinos do que a rede de operação com ELISA.

As Figuras 5.26(c) e 5.26(d) relacionam as saídas retornadas pelas redes de operação de Macaé para os dados da fazenda RS. Por suas saídas, foi possível notar um aumento na Sensibilidade desses animais quando ELISA foi adicionado aos testes cervicais, com a medida aumentando de 56% em *TCS TCC* para 68% em *TCS TCC ELISA*. Graficamente tal aumento é percebido com um deslocamento dos animais TB+ após o ponto de corte da rede de operação com ELISA. A medida de Especificidade permaneceu com 40% (2 animais) de performance mesmo com o uso de ELISA. Com o aumento visto na Sensibilidade, as medidas de Acuidade e SP também apresentaram mudanças com o uso de ELISA, aumentando 10 pontos percentuais na Acuidade e um aumento de 0,05 (10% do valor original) na medida SP.

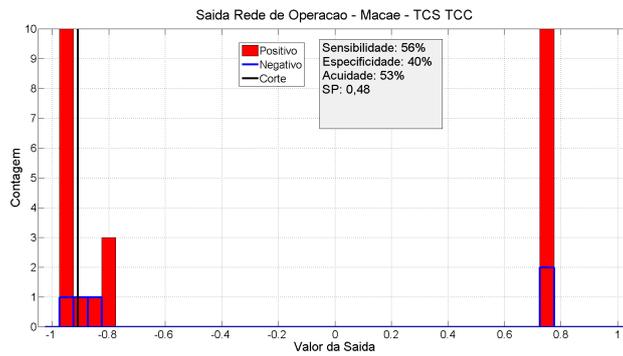
Observa-se ainda pelas figuras que as redes de operação de UFF apresentam mais saídas à direita para animais TB+ do que as redes de Macaé. Tal fato mostra que UFF aproxima melhor o seu resultado ao valor objetivo definido para um diagnóstico positivo de TB do que Macaé. No experimento anterior (Seção 5.1) também foi visto essa melhor performance na Sensibilidade de UFF ao comparar com os modelos de Macaé.



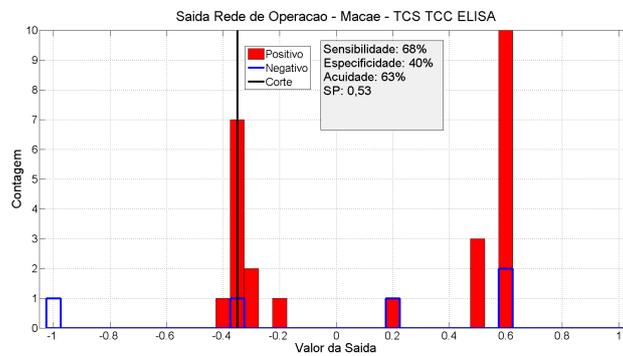
(a) Modelo TCS TCC de **UFF**



(b) Modelo TCS TCC ELISA de **UFF**



(c) Modelo TCS TCC de **Macaé**



(d) Modelo TCS TCC ELISA de **Macaé**

Figura 5.26: Saídas das Redes de Operação para os animais de RS, **COM** anérgicos no treinamento

A Tabela 5.23 relaciona as medidas de performance das propagações dos dados de RS nas redes de operação usadas, vistas na Figura 5.26.

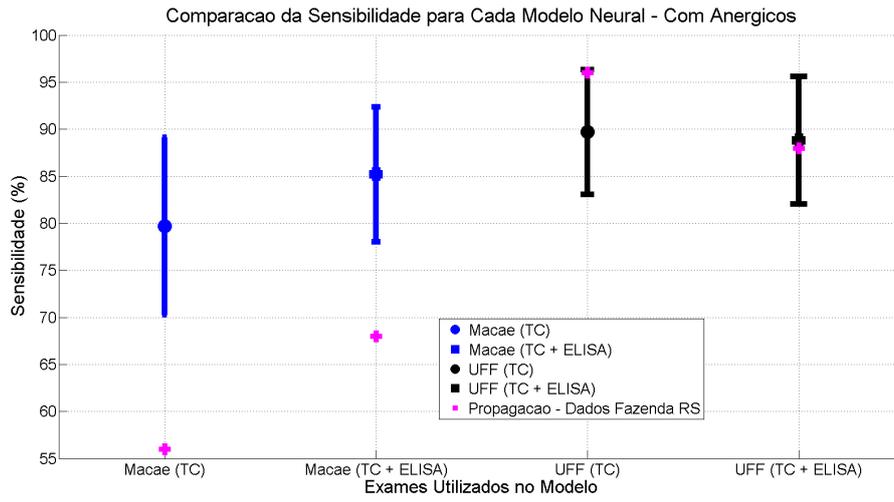
Tabela 5.23: Dados de RS propagados nas redes de operação treinadas COM anérgicos

Fazenda	Mod. Neural	Sens. (%)	Espec. (%)	Acc. (%)	SP
UFF	TCS TCC	96	20	83	0,50
UFF	TCS TCC ELISA	88	20	77	0,48
Macaé	TCS TCC	56	40	53	0,48
Macaé	TCS TCC ELISA	68	40	63	0,53

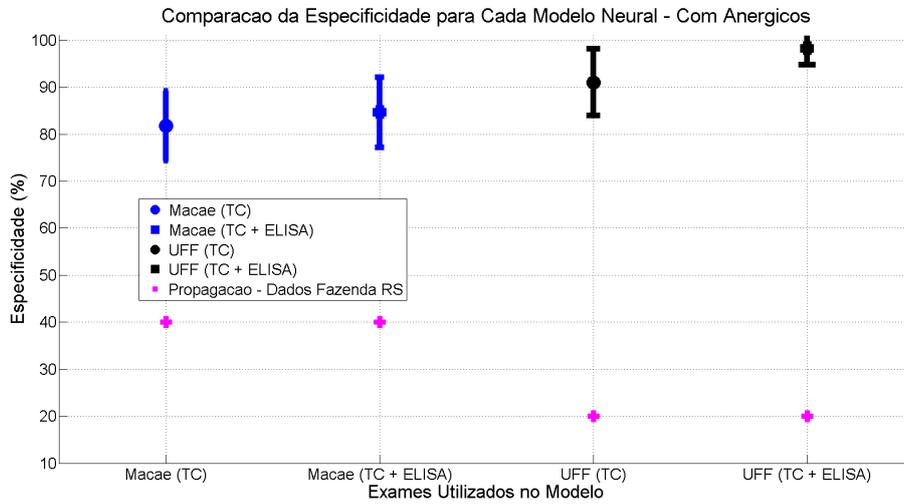
Comparando os modelos segundo a medida SP, viu-se que apenas nas redes de operação de Macaé houve aumento da medida com o acréscimo de ELISA, enquanto que para UFF houve uma diminuição. No entanto, elas apresentaram valores menores do que os valores médios observados nestes modelos (Seção 5.1), assim como o treinamento sem anérgicos.

Ainda para avaliar a classificação dos dados de RS pelas redes de operação consideradas neste estudo, verificou-se se as medidas de performance de RS se encontram dentro do erro estatístico dos modelos neurais de UFF e Macaé. As Tabelas 5.3 e 5.10 relacionam a variação de performance dos modelos de ambas fazendas. Assim como feito no caso de treinamento sem anérgicos, a medida SP foi desconsiderada nesta análise. A Figura 5.27 relaciona as performances dos dados de RS propagados nas redes de operação citadas somente para as medidas de Sensibilidade, Especificidade e Acuidade.

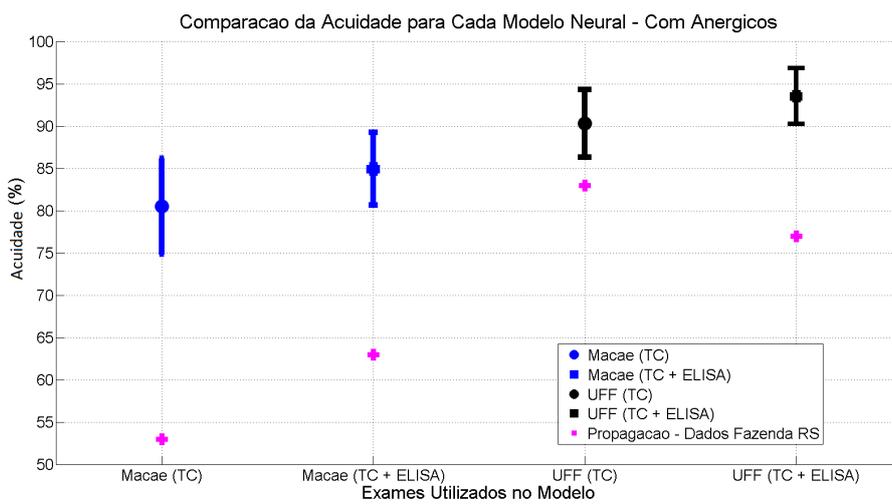
A comparação entre as performances das redes de operação e a variação de um modelo neural mostra novamente as mesmas características vistas para o caso sem anérgicos no treinamento. As medidas de Acuidade e Especificidade estavam fora das regiões de média e desvio padrão dos modelos de UFF e Macaé. A Acuidade em UFF apresentou-se próxima do comportamento dos modelos da fazenda, mas sua performance foi comprometida pela baixa Especificidade, além de haver queda nesta medida devido a Sensibilidade da rede de operação com ELISA. Em Macaé, sua Acuidade esteve distante dos intervalos de performance dos modelos devido à sua baixa taxa de Sensibilidade e Especificidade. Apenas a Sensibilidade de RS em UFF apresentou seus valores dentro do desvio padrão de seus modelos neurais, indicando que para animais TB+ o comportamento dos modelos de UFF se aplicam em RS.



(a) Sensibilidade



(b) Especificidade



(c) Acuidade

Figura 5.27: Dados da fazenda RS simulados nos modelos neurais de UFF e Macaé treinados **COM** anérgicos

Pelos resultados vistos neste experimento, pode-se concluir que os dados encontrados em RS possuem, ao menos com relação aos animais com Tuberculose positiva, semelhanças com a fazenda UFF, e que os efeitos com a adição de ELISA também podem ser vistos parcialmente nesta fazenda. A alta Sensibilidade gerada pelas redes de operação dos modelos neurais de UFF e o fato destes valores estarem dentro da variação estatística da fazenda, indica que a generalização de seus modelos foi suficiente para classificar corretamente os animais TB+ de RS.

Para a Especificidade, nenhuma das fazendas se mostrou capaz de classificar todos estes animais de acordo com diagnósticos que lhe foram definidos e esta medida pode ser explicada pelos dados da fazenda RS. O primeiro fator para a análise de Especificidade em RS é a pequena quantidade de animais definidos como TB-, havendo apenas 5 amostras na fazenda. Com uma quantidade pequena desses animais, a variação da Especificidade é alta, já que cada animal representa ganho ou perda de 20% na medida, dependendo do diagnóstico feito. Há ainda o fato dos dados estarem incompletos, faltando o exame confirmatório da presença ou não de *M. bovis* nos bovinos, o que impacta diretamente na performance de classificação dos modelos neurais.

Em todas as redes de operação testadas neste experimento, uma quantidade mínima de três animais TB- foram sistematicamente classificados como TB+ (Figuras 5.24 e 5.26). UFF foi a fazenda que apresentou as melhores medidas de Sensibilidade, e para este caso classificou sempre 4 destes animais como TB+, sendo alguns destes bovinos com resultados próximos de 1, o que mostra que a rede responde que a probabilidade deste animal ser um animal tuberculoso é maior. Logo, pelos resultados visto neste experimento, é possível que alguns dos animais definidos como negativos para Tuberculose em RS sejam na verdade animais doentes, e consequentemente a análise sobre as medidas de performance foi prejudicada em tal situação.

### 5.3 Otimização dos testes cervicais

Para o presente experimento, a proposta é determinar um novo ponto de corte para os testes cervicais, baseado no valor no SP. Em seguida, analisar os impactos que essa mudança gera na eficiência dos exames ao ser comparado com a regra hoje usada no Brasil. A análise é feita tanto sobre o TCS quando o TCC.

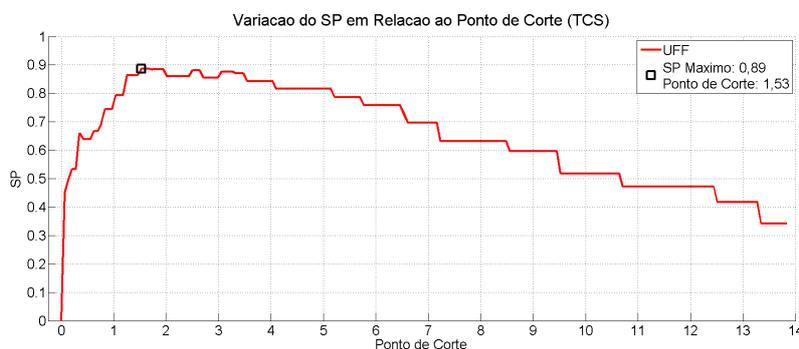
As eficiências nesta análise foram consideradas pelas medidas de Sensibilidade, Especificidade, Acuidade e valor SP. São analisados os cortes de cada exame tanto sobre os animais não anérgicos como também para os anérgicos de cada fazenda. Ainda é importante lembrar que neste experimento, um diagnóstico inconclusivo, pela norma do PNCEBT, foi considerado como um diagnóstico errado.

Na apresentação dos resultados, são citadas as “Região de diagnóstico” da norma preconizada pelo PNCEBT. A região chamada de “Negativa” refere-se ao intervalo onde um animal é considerado negativo para tuberculose ( $\Delta\text{Bov}$  ou  $\Delta\text{Bov} - \Delta\text{Av} < 2$ ), a região “Inconclusiva” para um diagnóstico inconclusivo ( $2 \leq \Delta\text{Bov}$  ou  $\Delta\text{Bov} - \Delta\text{Av} < 4$ ) e a região “Positiva” para quando um animal é dito como infectado pela doença ( $\Delta\text{Bov}$  ou  $\Delta\text{Bov} - \Delta\text{Av} \geq 4$ ).

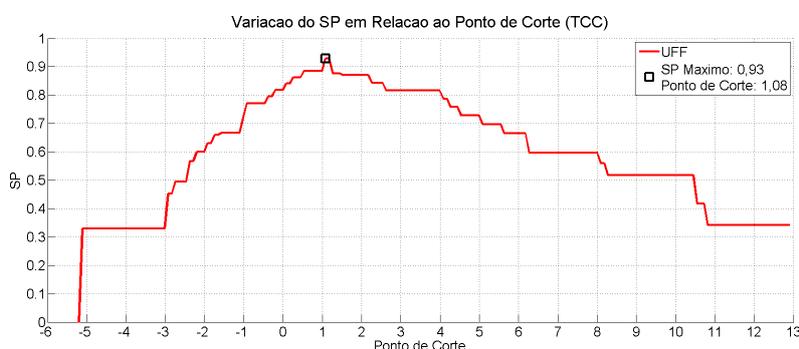
Cada fazenda tem seus resultados apresentados e discutidos separadamente, nas subseções 5.3.1 (UFF), 5.3.2 (Macaé), 5.3.3 (RS) e 5.3.4 em uma análise global de todas as fazendas.

### 5.3.1 Fazenda UFF

Para os dois testes cervicais de UFF, variou-se o ponto de corte dentro da faixa dinâmica dos valores de ( $\Delta\text{Bov}$ ) e ( $\Delta\text{Bov} - \Delta\text{Av}$ ), após o processamento dos dados descrito em 4.1.4, respectivamente para o TCS e o TCC. Para o exame TCS, sua faixa dinâmica varia de 0 até 13,90 para ( $\Delta\text{Bov}$ ), enquanto que para o exame TCC sua faixa dinâmica vai de  $-5,20$  até 13, para ( $\Delta\text{Bov} - \Delta\text{Av}$ ). A Figura 5.28 trás a variação do valor SP para os diferentes pontos de corte para os exames TCS e TCC.



(a) Curva para o exame TCS



(b) Curva para o exame TCC

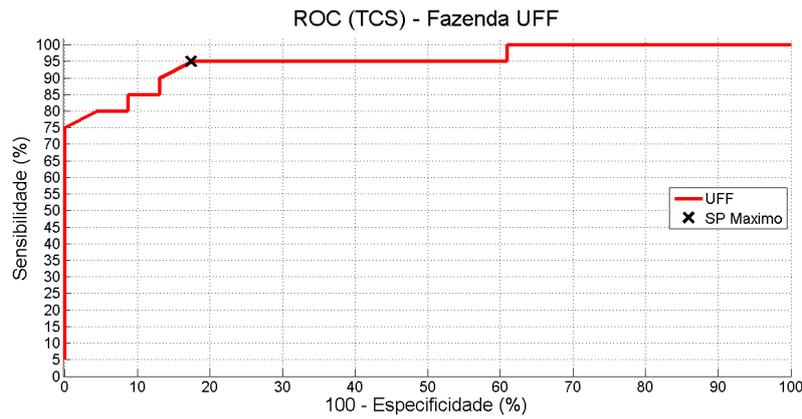
Figura 5.28: Variação do SP pelo ponto de corte da Fazenda UFF para os exames TCS e TCC

Considerando primeiro o teste cervical simples (TCS), o gráfico apresentado na Figura 5.28(a) mostra que dentre todos os pontos de corte testados, o ponto de corte de valor **1,53** foi o que apresentou o maior valor SP (0,89). Já para o exame TCC, a Figura 5.28(b) mostra que o ponto de corte de valor **1,08** apresentou um valor máximo SP de 0,93. A Tabela 5.24 trás a relação da faixa dinâmica dos exames, o novo ponto de corte e o valor SP associado à este novo corte escolhido.

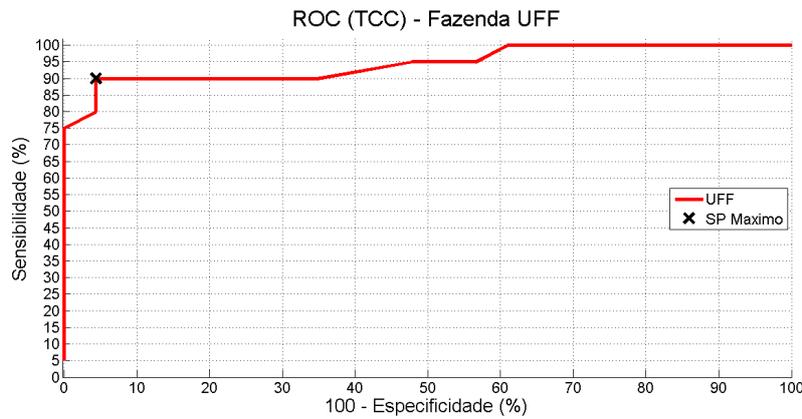
Tabela 5.24: Valores de Faixa dinâmica, Pontos de Corte e SP para TCS e TCC da fazenda UFF

<b>Exame</b>	<b>Faixa Dinâmica</b>	<b>Ponto de Corte</b>	<b>SP</b>
TCS	[0 ; 13,90]	1,53	0,89
TCC	[-5,20 ; 13]	1,08	0,93

Na Figura 5.28, tanto para o TCS quanto para o TCC, o valor SP aumenta a medida que o ponto de corte testado se movimentava para a direita, encontrando um valor máximo para depois apresentar uma queda até o fim do intervalo considerado. Uma vez que o valor SP busca equilibrar a Sensibilidade e a Especificidade, tal queda no valor SP quer dizer que uma das medidas apresenta queda ao passo que aumenta o valor do ponto de corte. A Figura 5.29 apresenta as curvas ROC para o TCS e o TCC de UFF, com a marcação do ponto de corte que apresentou o valor máximo de SP e através destas é possível identificar para os exames, as medidas de performance de Sensibilidade, Especificidade, Acuidade. Em sua curva, o TCS apresenta no ponto de máximo SP uma Sensibilidade de 95% e uma Especificidade de 83%, e o TCC apresenta uma Sensibilidade de 90% e uma Especificidade de 96% para o ponto de maior valor SP.



(a) Curva ROC do exame TCS



(b) Curva ROC do exame TCC

Figura 5.29: Curvas ROC da Fazenda UFF para os exames TCS e TCC

O passo seguinte nos resultados é dado pela comparação do ponto de corte baseado no valor SP com os pontos de cortes definidos pelo PNCEBT. A Tabela 5.25 relaciona as medidas de performance consideradas neste estudo, para os cortes oficiais e para os novos cortes de TCS e TCC, somente para animais não anérgicos. Para a norma oficial, suas medidas de avaliação podem ser calculadas a partir dos dados disponibilizados (Seção 4.1.1), enquanto que para o corte por SP suas medidas são geradas pelas curvas de variação do SP e ROC.

Tabela 5.25: Eficiência das normas para exames cervicais para a fazenda UFF - Animais Não Anérgicos

<b>Exame</b>	<b>Norma</b>	<b>Sens. (%)</b>	<b>Espec. (%)</b>	<b>Acc. (%)</b>	<b>SP</b>
TCS	PNCEBT	70	87	79	0,78
TCS	SP	95	83	88	0,89
TCC	PNCEBT	65	100	84	0,82
TCC	SP	90	96	93	0,93

Os exames cervicais com os cortes otimizados apresentaram um aumento de 0,11 (14%) no SP, consideravelmente alto perante a norma brasileira. Houve também um aumento da Sensibilidade para cada exame ao adotar o corte otimizado, onde o exame TCS aumentou de 70% de para 95%, enquanto o exame TCC aumentou de 65% para 90% na Sensibilidade. A otimização, no entanto, gerou perdas na Especificidade dos exames, com o TCS diminuindo de 87% para 83% no TCS, e de 100% para 96% no TCC. Os resultados mostram que enquanto há um ganho na Sensibilidade também há uma perda na Especificidade dos testes cervicais, indicando que existe uma área de confusão para estes exames onde há animais negativos sendo classificados como positivos para TB. No entanto, tais mudanças em Sensibilidade e Especificidade não foram observadas na mesma intensidade e proporção, indicando um pequeno aumento na quantidade de diagnóstico de falsos positivos por parte dos novos cortes.

A Figura 5.30 é usada de maneira complementar à análise da Tabela 5.25, e apresenta a distribuição dos animais não anérgicos de UFF para os testes cervicais simples e comparativo. A figura mostra que tanto para o TCS quanto para o TCC, os cortes baseado em SP estavam sistematicamente abaixo do corte determinado pelo PNCEBT para o diagnóstico negativo de TB. O principal fator para que os cortes estivessem abaixo da norma brasileira foi a distribuição dos animais tuberculosos e não tuberculosos pelos exames dentro de suas faixas dinâmicas.

Considerando primeiramente os animais TB−, para os dois testes cervicais observou-se que estes animais estavam quase que totalmente concentrados abaixo do corte de diagnóstico Negativo da norma brasileira, apresentando apenas alguns animais na região de Inconclusivo. Em contrapartida, os animais TB+ estavam dispersos entre todas as regiões de diagnóstico, estando mais presentes na região de diagnóstico Positivo e Inconclusivo da norma do PNCEBT ( $\Delta_{Bov} \geq 2$  &  $\Delta_{Bov} - \Delta_{Av} \geq 2$ ). Os animais TB+ apareceram em menor quantidade na região de diagnóstico Negativo, porém próximos do corte para região de Inconclusivo. Já que o novo corte é baseado no valor SP e esta medida busca maximizar e equilibrar a

Sensibilidade e Especificidade, este foi colocado na região limite em que acaba a concentração de animais TB- e que começam a ocorrer os casos de TB+, porém ainda apresentando falsos negativos e positivos.

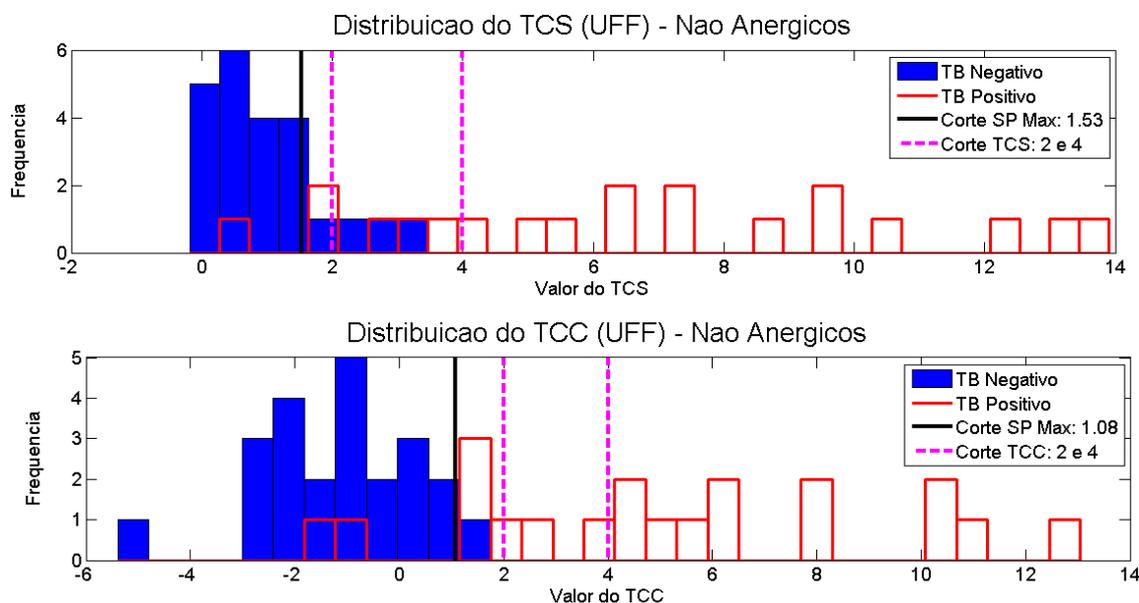


Figura 5.30: Testes Cervicais de UFF para animais não anérgicos

A análise também é estendida para os animais anérgicos da fazenda. A Tabela 5.26 trás a Sensibilidade do corte oficial e do novo corte para o TCS e o TCC aplicados à esses animais. O uso do ponto de corte otimizado apresentou um ganho na Sensibilidade dos anérgicos em ambos os testes, fazendo que para o TCS houvesse um aumento de 33% para 100% e no TCC houvesse aumento de 0% para 67% na Sensibilidade. Os ganhos na porcentagem de identificação são significativos devido a pequena quantidade de animais anérgicos presentes na fazenda (3 animais).

Tabela 5.26: Eficiência das normas para cervicais para a fazenda UFF - Animais Anérgicos

Exame	Norma	Sensibilidade (%)
TCS	PNCEBT	33
TCS	SP	100
TCC	PNCEBT	0
TCC	SP	67

A suspeita de anergia em bovinos é considerada quando um animal infectado por *M. bovis* não apresenta reação aos testes cervicais ao ponto de indicarem a presença do agente. Em UFF, os animais anérgicos foram assim considerados pela sua não

reação apenas ao TCC e uma posterior identificação por parte de ELISA. A Tabela 5.26 esclarece tal fato, mostrando que o TCC, segundo a norma oficial, não identificou nenhum animal, enquanto o TCS conseguiu ainda identificar corretamente um animal anérgico.

A Figura 5.31 apresenta a distribuição dos animais anérgicos de UFF em ambos os testes, e os pontos de corte analisados neste experimento.

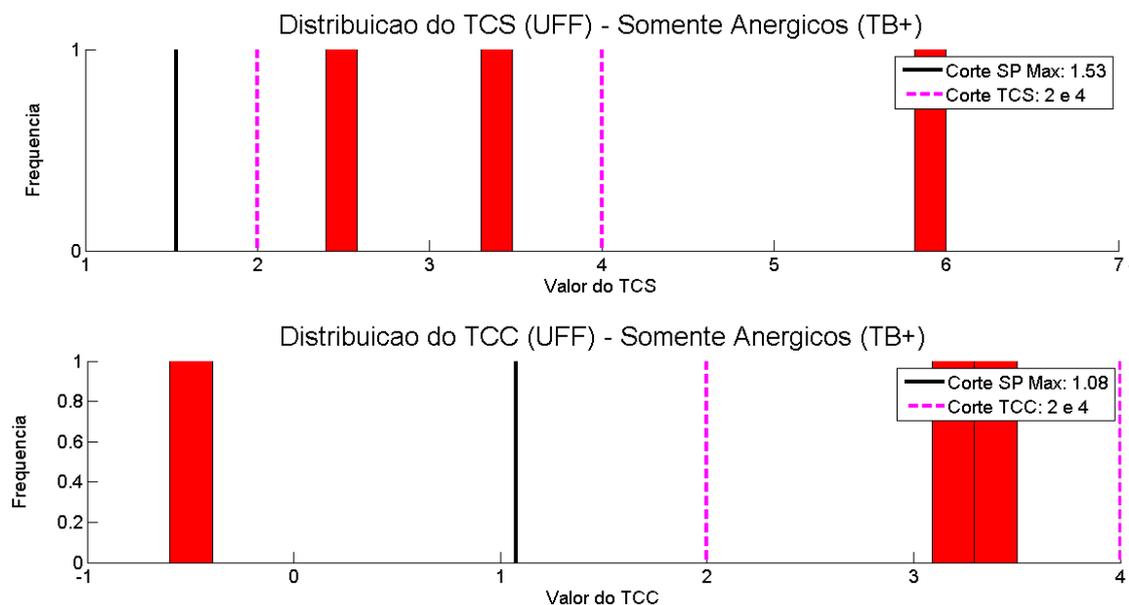


Figura 5.31: Testes Cervicais de UFF para animais anérgicos

Analisando o TCS, dois animais apresentaram seus resultados na zona de Inconclusivo ( $2 \leq \Delta\text{Bov} < 4$ ) e um animal na região de diagnóstico Positivo ( $\Delta\text{Bov} \geq 4$ ) da norma oficial. Com o novo corte definido com um valor inferior ao limite da região de diagnóstico Negativo ( $\Delta\text{Bov} < 2$ ), o TCS consegue identificar todos os animais anérgicos da fazenda. Para o TCC, o que se observa é que apenas um animal se encontra na região de diagnóstico Negativo da norma oficial ( $\Delta\text{Bov} - \Delta\text{Av} < 2$ ) com um valor distante ao corte igual a 2, e dois animais na região de “Inconclusivo” ( $2 \leq \Delta\text{Bov} - \Delta\text{Av} < 4$ ). O corte otimizado para o TCC consegue classificar corretamente apenas os animais da região de “Inconclusivos”, porém não foi capaz de diagnosticar o terceiro animal anérgico. Um ponto de corte capaz de classificar corretamente todos os animais anérgicos, acarretaria em uma perda na Especificidade para os animais não anérgicos já que parte destes animais não seriam corretamente classificados como TB-.

Os novos pontos de corte em UFF apresentaram ganhos na eficiência para os seus testes cervicais quando comparados com os cortes definidos pelo PNCEBT. Tanto para os animais anérgicos como para os não anérgicos, o novo corte conseguiu classificar animais que pela norma oficial não eram classificados corretamente. A principal limitação observada para os cortes oficiais foi a classificação de animais

ditos como inconclusivos, já que tal diagnóstico é tido como incorreto para a presença ou ausência de TB. Os novos cortes, conseguiram classificar parte dos animais presentes nesta região de diagnóstico Inconclusivo e conseqüentemente, aumentando a eficiência dos testes para suas Sensibilidade e Especificidade.

### 5.3.2 Fazenda Macaé

Na fazenda Macaé, diferentes pontos de corte foram testados nas respectivas faixas dinâmicas de seus testes cervicais, sendo este intervalo de variação determinado à partir das tuberculinas da fazenda, assim como feito para UFF (Seção 5.3.1). Para o TCS, o valor de tuberculina bovina ( $\Delta\text{Bov}$ ) varia de 0,20 mm até no máximo 19,40 mm, enquanto que para o TCC a diferença entre as tuberculinas bovina e aviária ( $\Delta\text{Bov} - \Delta\text{Av}$ ) varia entre os valores  $-2,40$  mm até 16,70 mm. A Figura 5.32 mostra variação do SP para os pontos de corte dos testes cervicais simples e comparativo da fazenda Macaé e vê-se que para o TCS (Figura 5.32(a)) que seu ponto de corte ótimo é definido em **4,81**. O valor SP associado à tal corte é igual a 0,82 e este é o valor máximo dentre todos os valores SP encontrados. Já para o TCC (Figura 5.32(b)) é mostrado que o ponto de corte otimizado apresentou um valor máximo de SP igual à 0,84 e este valor foi encontrado para o ponto de corte de valor **1,80**. Na Tabela 5.27 são expostos as faixas dinâmicas de ambos os testes cervicais, o ponto de corte otimizado encontrado para cada um, e por fim, o valor SP associado ao novo ponto de corte gerado.

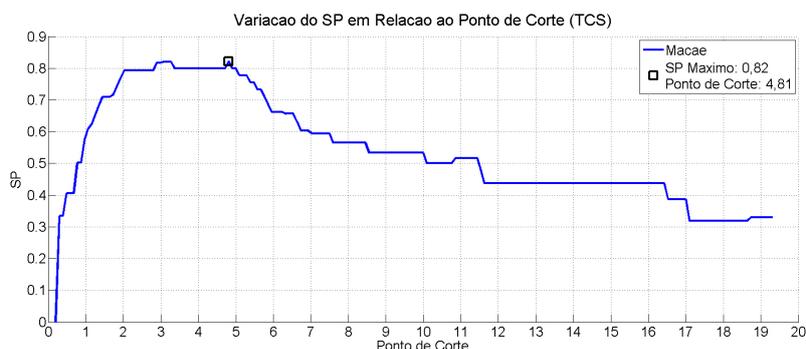
Tabela 5.27: Valores de Faixa dinâmica, Pontos de Corte e SP para TCS e TCC da fazenda Macaé

Exame	Faixa Dinâmica	Ponto de Corte	SP
TCS	[0,2 ; 19,40]	4,81	0,82
TCC	[-2,40 ; 16,70]	1,80	0,84

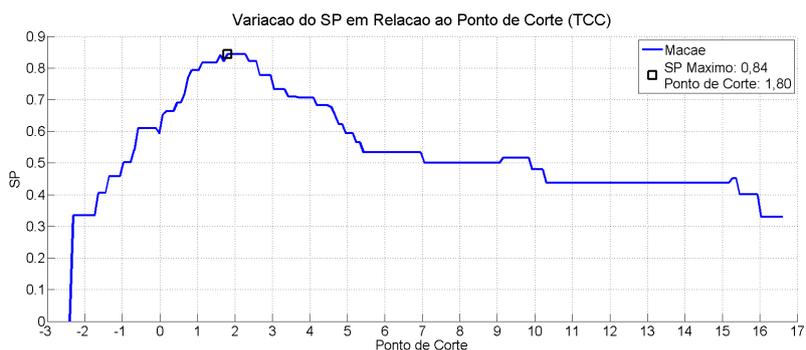
Uma diferença importante foi observada entre o TCS e o TCC de Macaé, onde o novo corte do TCC (**1,80**) esteve abaixo do menor corte definido pelo PNCEBT, igual a 2, o que também foi observado para a fazenda UFF. O corte para o TCS (**4,81**) foi definido acima do corte que define que um animal está com tuberculose preconizado pelo PNCEBT, ou seja, superior a 4. Tal ponto de corte de TCS indica que há uma baixa Especificidade para o corte oficial, uma vez que quanto mais a direita estiver um corte, maior é a sua Especificidade e podendo também acarretar em queda na Sensibilidade, dependendo da amostra analisada. Já para o TCC, um ponto de corte mais próximo à esquerda indica um aumento na Sensibilidade,

podendo este gerar uma diminuição da Especificidade, novamente dependendo dos dados analisados.

O comportamento apresentado no gráfico de variação do SP pelos pontos de cortes de UFF e Macaé (Figuras 5.28 e 5.32) apresentam semelhanças, como o aumento do valor SP a medida que o ponto de corte testado se movimenta para a direita e uma queda até o fim da faixa dinâmica do exame após o pico de máximo SP.



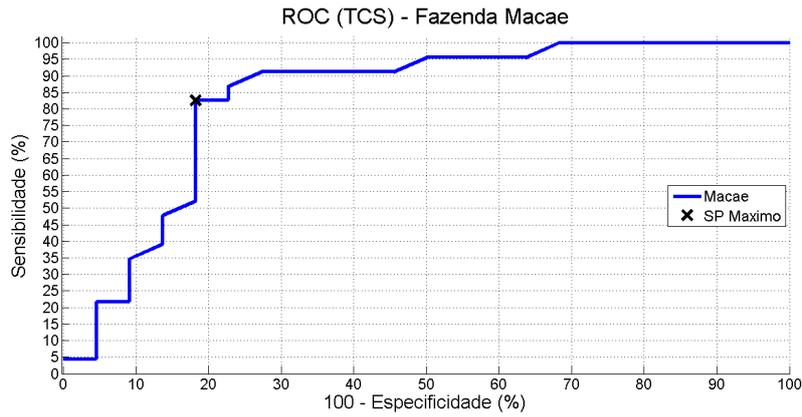
(a) Curva para o exame TCS



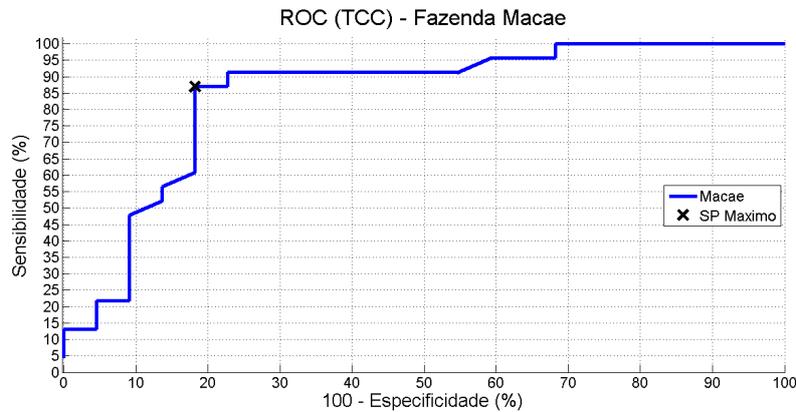
(b) Curva para o exame TCC

Figura 5.32: Variação do SP pelo ponto de corte da Fazenda Macaé para os exames TCS e TCC

A análise por curva ROC é usada para determinar as medidas de performance de Sensibilidade, Especificidade e Acuidade nos pontos de corte gerados neste experimento. A Figura 5.33 apresenta as curvas ROC para o TCS e o TCC de Macaé, destacando o ponto de corte com valor máximo de SP. Na curva ROC do TCS (Figura 5.33(a)), o ponto de corte otimizado apresenta uma Sensibilidade de 83% e uma Especificidade de 82%, enquanto que para o TCC (Figura 5.33(b)) sua curva indica uma Sensibilidade de 87% e uma Especificidade de 82% para o seu ponto de corte ótimo. A Acuidade foi calculada considerando a quantidade de animais que cada porcentagem de Sensibilidade e Especificidade representa, sendo encontrado uma Acuidade de 82% para o TCS e 84% para o TCC.



(a) Curva ROC do exame TCS



(b) Curva ROC do exame TCC

Figura 5.33: Curvas ROC da Fazenda Macaé para os exames TCS e TCC

A comparação entre as duas normas de corte é feita primeiramente considerando apenas os animais não anérgicos de Macaé, com suas medidas de performance apresentadas na Tabela 5.28. As medidas de Sensibilidade, Especificidade e Acuidade dos cortes definidos pelo PNCEBT foram calculados a partir dos dados disponíveis da fazenda (Seção 4.1.2). No TCS, observou-se um aumento de 64% na Especificidade dos cortes oficiais para 82% no novo ponto de corte, enquanto a Sensibilidade não apresentou alteração e se manteve com uma porcentagem de 83%. Tal aumento na Especificidade acarretou em um aumento de 9% para a Acuidade do teste e também em seu valor SP, passando de 0,73 para 0,82. Para o TCC ocorreu a situação contrária à ocorrida ao teste cervical simples, tendo a Especificidade se mantido constante (82%) para ambos os cortes considerados e a Sensibilidade aumentando de 57% para 87% em eficiência quando o novo corte é usado. Com este aumento na Sensibilidade, a Acuidade do TCC aumentou de 69% para 84% como também sua medida de SP, passando de 0,69 para 0,84.

Tabela 5.28: Eficiência das normas para exames cervicais para a fazenda Macaé - Animais Não Anérgicos

<b>Exame</b>	<b>Norma</b>	<b>Sens. (%)</b>	<b>Espec. (%)</b>	<b>Acc. (%)</b>	<b>SP</b>
TCS	PNCEBT	83	64	73	0,73
TCS	SP	83	82	82	0,82
TCC	PNCEBT	57	82	69	0,69
TCC	SP	87	82	84	0,84

A Figura 5.34 é usada de maneira complementar às análises dos resultados da Tabela 5.28, mostrando a aplicação do ponto corte por SP e dos pontos de corte preconizados pela norma brasileira, tal como a distribuição dos animais não anérgicos de Macaé. Na distribuição dos animais para o TCS, viu-se uma dispersão de 7 animais TB– nas regiões de diagnóstico “Inconclusivo” e “Positivo” da norma oficial, representando um terço do total deste tipo de animal na fazenda. Para os animais TB+ apenas 4 animais, de um total de 23, estão fora da região correta para sua classificação, aparecendo nas regiões de diagnóstico “Negativo” e “Inconclusivo”. O novo ponto de corte reduziu a classificação incorreta de animais TB– para 4 animais e manteve também em 4 animais o erro de classificação dos animais TB+, o que gerou um aumento da Especificidade e do SP em TCS. Logo, a razão para seu ponto de corte elevado foi a distribuição dos animais sadios de Macaé.

Na distribuição dos animais para o TCC foi observado uma concentração de animais TB+ na região de diagnóstico “Inconclusivo” da norma preconizada pelo PNCEBT. Neste caso, um ponto de corte superior a 4 e similar ao corte ótimo do TCS apresentaria uma queda na Sensibilidade do exame, já que este não conseguiria classificar corretamente os animais tuberculosos desta região da faixa dinâmica. Com o novo corte definido próximo ao corte oficial igual a 2, o novo corte reduziu o diagnóstico de falsos negativos de 10 para apenas 3 animais, o que explica o aumento de 30% da Sensibilidade do exame. Para os animais TB–, 3 animais eram ditos como positivos e 1 animal dito como inconclusivo para a norma oficial, e pelo novo corte estes mesmos animais passaram a ser classificados como falsos positivos, o que reflete na não alteração do valor de Especificidade do TCC em 82%. Logo, para o TCC a definição de seu corte em **1,80** foi decidida pela concentração de animais tuberculosos na região de diagnóstico Inconclusivo e pela concentração de animais TB– também próximos à essa mesma região.

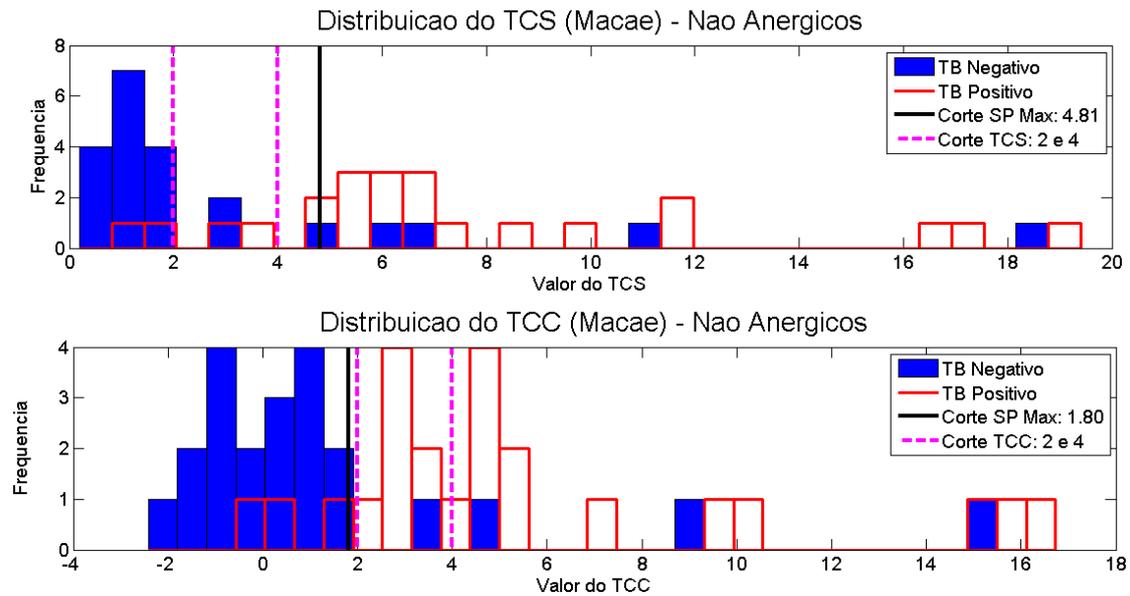


Figura 5.34: Testes Cervicais de Macaé para animais não anérgicos

A Tabela 5.29 trata do diagnóstico dos animais anérgicos de Macaé pelos pontos de corte estudados. Para nenhuma das metodologias houve ganho ou perda na classificação de tais animais. Nenhum dos cortes conseguiu classificar qualquer um dos quatro anérgicos presentes na fazenda, para nenhum dos dois exames, obtendo-se para todos os casos uma Sensibilidade igual a 0%.

Tabela 5.29: Eficiência das normas para exames cervicais para a fazenda Macaé - animais anérgicos

Exame	Norma	Sensibilidade (%)
TCS	PNCEBT	0
TCS	SP	0
TCC	PNCEBT	0
TCC	SP	0

A Figura 5.35 apresenta a distribuição dos animais anérgicos analisados e os pontos de corte considerados para TCS e TCC. Todos os animais anérgicos de Macaé tiveram seus diagnósticos definidos como negativos para o TCC pelos cortes preconizados pelo PNCEBT, enquanto para o TCS dois animais tiveram seus diagnósticos definidos como negativo e os demais ditos como inconclusivo pela norma oficial. Os pontos de corte otimizados por SP foram determinados levando-se em conta todo o conjunto de animais da fazenda e assim, ambos se apresentaram acima dos valores dos animais anérgicos para cada teste cervical. No TCC o ponto de corte otimizado (1,80) se mostrou próximo à um anérgico ( $\Delta_{Bov} - \Delta_{Av}$  igual à 1,70), porém ainda

incapaz de classificá-lo corretamente.

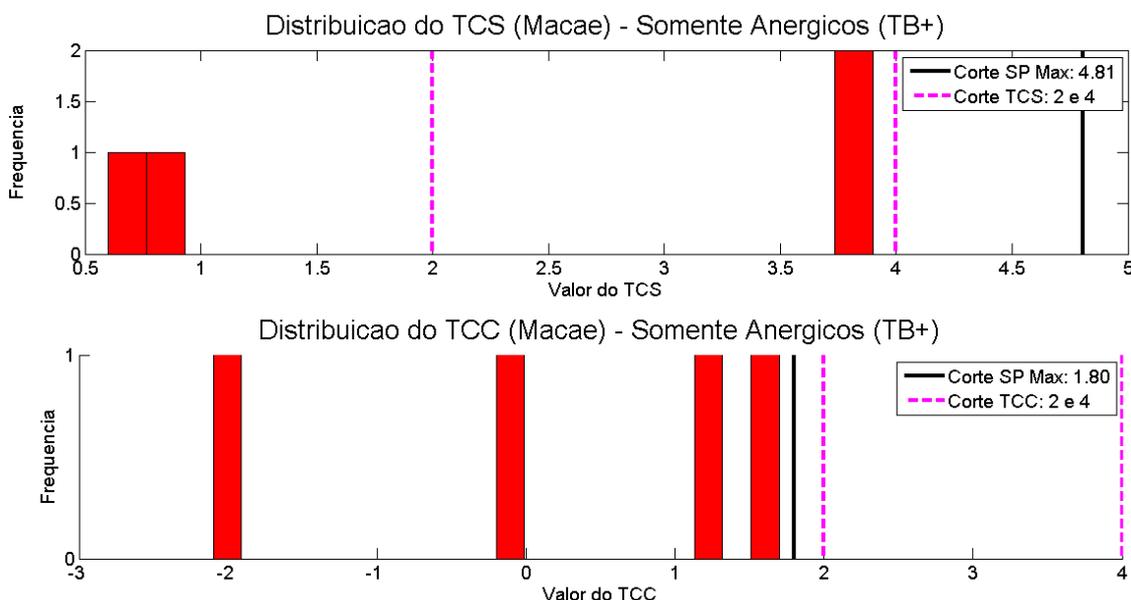


Figura 5.35: Testes Cervicais de Macaé para animais anérgicos

O método otimizado para os testes cervicais apresentou ganhos na eficiência dos exames, aumentando a Especificidade para o TCC e a Sensibilidade para o TCS e se mostrando vantajosa à fazenda Macaé. A fazenda apresentou uma maior incidência de falsos positivos, quando comparada à fazenda UFF, porém o maior gargalo visto para os cortes preconizados pela norma brasileira foi a região de animais inconclusivos. O TCC foi o teste que apresentou a maior concentração de animais nessa região de diagnóstico, com 7 animais TB+ que eram classificados erradamente pelos cortes oficiais.

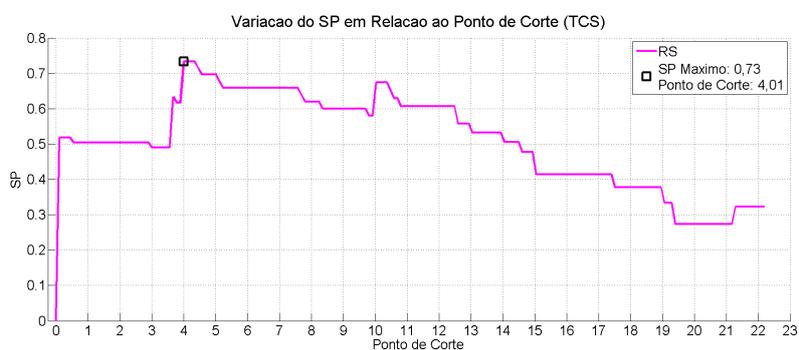
### 5.3.3 Fazenda RS

Na análise de RS, os resultados foram gerados considerando-se todos os animais como sendo não anérgicos, uma vez que os dados da fazenda não possuem a indicação para este tipo de bovino. A faixa dinâmica do TCS em RS varia de 0 até 22, 30 mm para a tuberculina bovina. No TCC, a variação da diferença entre as tuberculinas bovina e aviária varia de  $-4,90$  até 15. Nestas faixas dinâmicas são gerados e distribuídos uniformemente 200 pontos de corte para cada teste cervical, e a Figura 5.36 apresenta a variação do valor SP para cada corte testado e também o corte associado ao maior valor SP encontrado. Na curva do TCS (Figura 5.36(a)) foi encontrado o valor máximo de SP igual a 0,73 para esta medida, e o ponto de corte que gerou este valor para a medida SP foi **4,01**. Já para o TCC (Figura 5.36(b)) é mostrado que o ponto de corte otimizado apresentou um valor máximo de SP igual a 0,54, com o ponto de corte associado igual a **2,56**. Na Tabela 5.30 são expostas

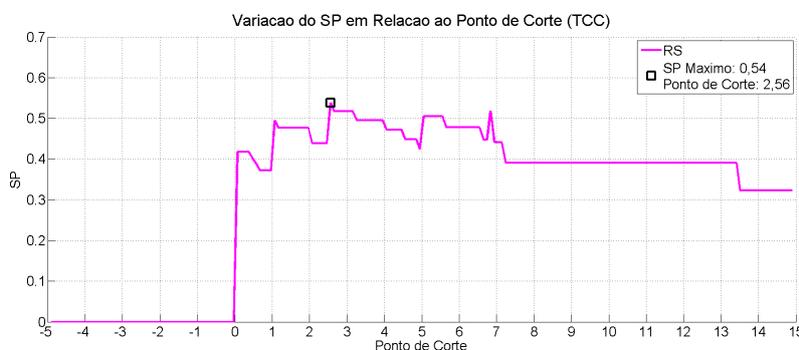
as faixas dinâmicas de ambos os testes cervicais, os pontos de corte de melhor SP e o valor desta medida associado ao corte.

Tabela 5.30: Valores de Faixa dinâmica, Pontos de Corte e SP para TCS e TCC da fazenda RS

Exame	Faixa Dinâmica	Ponto de Corte	SP
TCS	[0 ; 22,30]	4,01	0,73
TCC	[-4,90 ; 15]	2,56	0,54



(a) Curva para o exame TCS

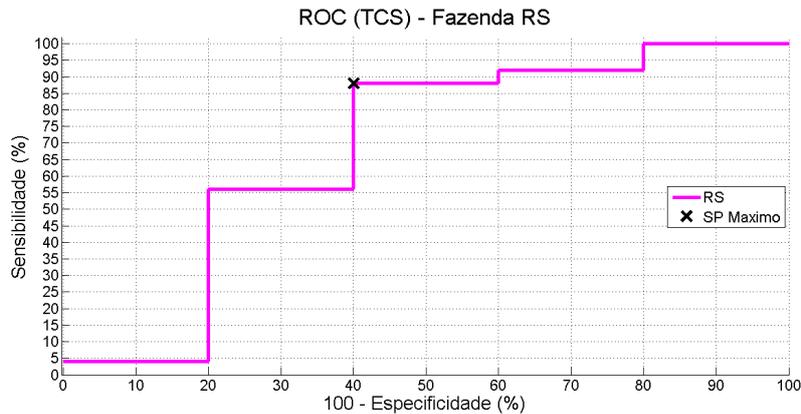


(b) Curva para o exame TCC

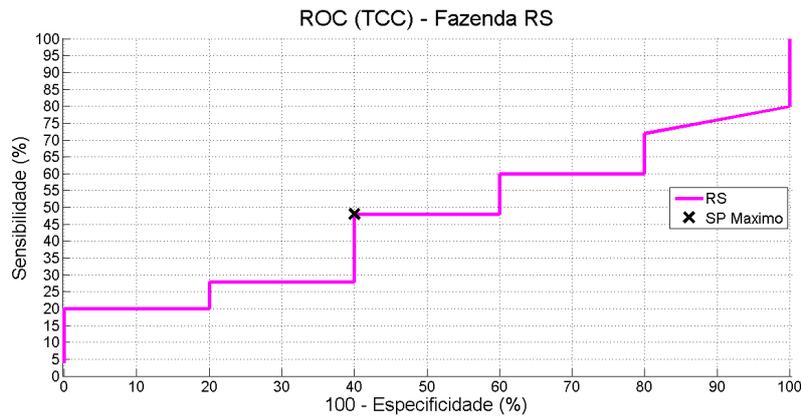
Figura 5.36: Variação do SP pelo ponto de corte da Fazenda RS para os exames TCS e TCC

O passo seguinte para os resultados foi determinar os valores das medidas de Sensibilidade, Especificidade e Acuidade para cada um dos novos pontos gerados. Para isso foram utilizadas as curvas ROC (Figura 5.37) geradas para os pontos de corte avaliados nos testes cervicais. Em cada curva ROC, foram observados os valores de Sensibilidade e Especificidade de um dado ponto de corte e na mesma curva é destacado o ponto com maior SP. A Figura 5.37(a) representa a curva ROC para o TCS de RS, e nela o ponto de corte otimizado apresenta uma Sensibilidade

de 88% e uma Especificidade de 60%. A Figura 5.37(b) apresenta a curva ROC para o TCC, com seu ponto de corte otimizado apresentando uma Sensibilidade de 60% e uma Especificidade de 50%. Com os valores das duas primeiras medidas e a quantidade de animais presentes na fazenda (Tabela 4.8), foram obtidos os valores de Acuidade para cada teste, com o TCS apresentando 83% e o TCC 50% com os novos cortes propostos.



(a) Curva ROC do exame TCS



(b) Curva ROC do exame TCC

Figura 5.37: Curvas ROC da Fazenda RS para os exames TCS e TCC

A Tabela 5.31 contém a relação das medidas de performance consideradas neste estudo para ambos os testes cervicais. A tabela relaciona cada teste intradérmico por seu ponto de corte baseado em SP e pelos cortes preconizados pela norma brasileira, somente para animais não anérgicos. Para os cortes oficiais, todas as medidas de performance consideradas foram calculadas à partir dos dados da fazenda (Seção 4.1.3), enquanto que para os novos cortes as suas performances foram obtidas através de suas curvas ROC e variação de SP por ponto de corte.

Considerando o valor SP, os cortes otimizados apresentaram um aumento para esta medida perante a norma brasileira, com o SP aumentando de 0,48 até 0,73

para o TCS e o TCC apresentando um aumento de 0,40 até 0,54. Houve também um aumento da Especificidade para cada exame ao adotar o corte otimizado, com o exame TCS aumentando de 20% de para 60% e o exame TCC aumentando de 40% para 60%. A Sensibilidade para o TCS não apresentou alteração, mantendo-se com 88% de eficiência, mas tal medida aumentou de 40% para 48% no TCC. Como consequência ao aumento das medidas de Especificidade e Sensibilidade, a Acuidade dos testes também apresentou ganhos. A Acuidade do TCS variou seu valor de 77% para 83% com o uso do novo ponto de corte, enquanto o TCC apresentou um ganho de 10%, passando de 40% pela norma preconizada pelo PNCEBT para 50% com o novo corte.

Tabela 5.31: Eficiência das normas para exames cervicais para a fazenda RS - animais não anérgicos

<b>Exame</b>	<b>Norma</b>	<b>Sens. (%)</b>	<b>Espec. (%)</b>	<b>Acc. (%)</b>	<b>SP</b>
TCS	PNCEBT	88	20	77	0,48
TCS	SP	88	60	83	0,73
TCC	PNCEBT	40	40	40	0,40
TCC	SP	48	60	50	0,54

Ainda comparando os cortes oficiais com os novos cortes otimizados, a Figura 5.38 apresenta a distribuição dos animais da fazenda para cada um dos testes cervicais, explicitando os pontos de corte de cada norma. Considerando inicialmente o TCS, o ponto de corte otimizado por SP apresentou uma diferença de apenas 0,01 com relação ao corte oficial que define a região de diagnóstico Positiva. Pela diferença nos cortes ser muito pequena, ambos adotam o mesmo critério para o diagnóstico de animais TB+, e conseqüentemente apresentam a mesma performance de Sensibilidade. No entanto, quanto a Especificidade a norma brasileira apresenta baixa performance uma vez que ela considera o diagnóstico “Inconclusivo” para os bovinos e, para o caso específico de RS, existem dois animais TB– presentes nessa região que passam a ser corretamente classificados pelo corte otimizado.

Para o TCC, a Figura 5.38 mostra que há uma concentração de animais TB+ com valores menores que os pontos de corte tanto para a norma oficial como para os cortes baseados em SP. Tal fato explica a baixa Sensibilidade do TCC em ambos os cortes, uma vez que para os cortes oficiais seus valores já estão fixados e o corte baseado em SP busca atingir um valor de Especificidade que ainda consiga maximizar a medida SP. Novamente o que se observa é que os cortes oficiais preconizados pela norma brasileira falham em classificar animais presentes em sua região de “Inconclusivo” e que o novo corte proposto consegue um aumento de 8% na Sensibilidade ao passar

a classificar estes animais.

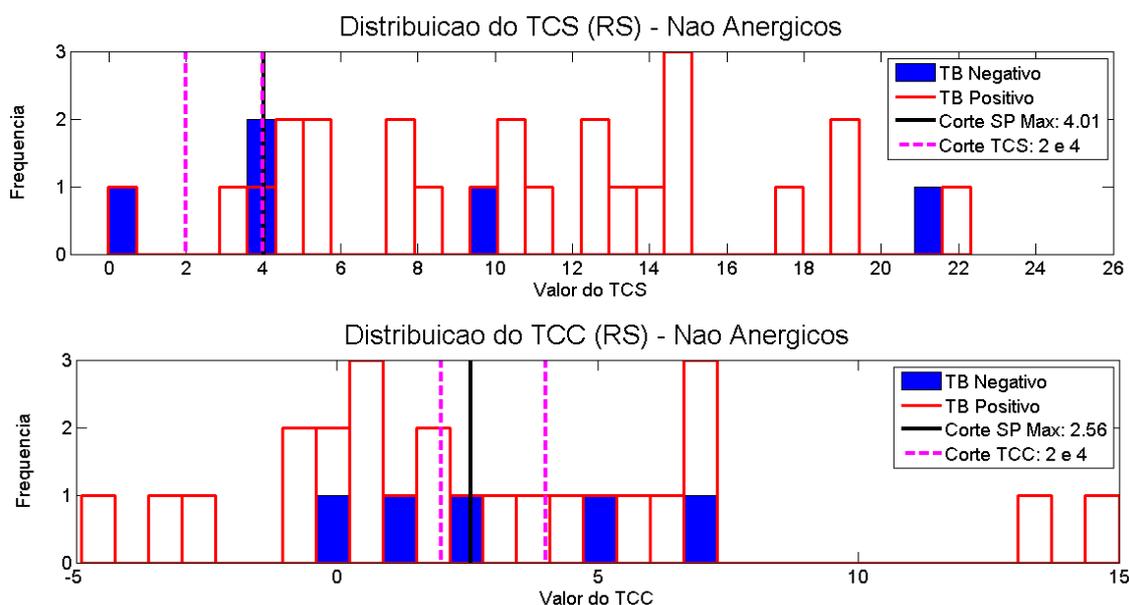


Figura 5.38: Testes Cervicais de RS para animais não anérgicos

Outro fator relevante para o TCS e o TCC de RS é a pequena quantidade de animais TB– da fazenda. Assim, cada animal classificado corretamente representa uma grande porcentagem para a Especificidade (20% por animal) e consequentemente para a medida SP. Tal característica foi vista nas curvas de variação de SP pelos pontos de corte (Figura 5.36), onde aumentos ou quedas acentuadas no valor SP indicam grandes mudanças na medida de Sensibilidade ou Especificidade. Na curva do TCS, imediatamente antes do pico de máximo SP, viu-se um aumento acentuado da curva, e ao analisar a Figura 5.38 percebeu-se que o novo corte consegue classificar corretamente dois animais TB–, o que gera um impacto de mais 40% na Especificidade do teste. Em TCC tal fato também ocorre, visto que à partir do início da faixa dinâmica do teste até o ponto de máximo de SP houve três aumentos abruptos do valor SP. Tais aumentos ocorrem quando os pontos de corte que passam a classificar corretamente cada um dos três animais TB– presentes, até o corte ótimo. Após cada aumento de SP foram vistas quedas graduais no mesmo, indicando que a medida que novos cortes são testados, ocorreram perdas na Sensibilidade. Tal queda na Sensibilidade é esperada, haja visto que na região anterior ao ponto de corte ótimo há uma concentração de animais TB+ e quanto mais à direita se mover o corte, menor será tal medida.

A fazenda RS apresentou melhores resultados na identificação da tuberculose com o uso do método otimizado para os testes cervicais. Nesta fazenda, a coleta dos dados ainda está em andamento e alguns dos animais considerados como livres de tuberculose podem ser na realidade animais infectados, porém novamente foi observado, principalmente para o TCC, que a norma oficial perde capacidade de

diagnóstico ao considerar um diagnóstico inconclusivo. A vantagem do novo corte proposto está em desconsiderar o diagnóstico inconclusivo da norma oficial.

### 5.3.4 Fazenda All

A fazenda *All* é uma fazenda fictícia considerada para uma abordagem global dos testes cervicais. Seus dados são compostos pelos animais das fazendas UFF, Macaé e RS, com seus animais anérgicos e não anérgicos. Com esta junção dos dados, a fazenda acabou composta por um total de 125 animais, sendo 68 animais tuberculosos, 50 animais livres da doença e 7 anérgicos. A Tabela 5.32 relaciona esta divisão dos bovinos na fazenda *All*.

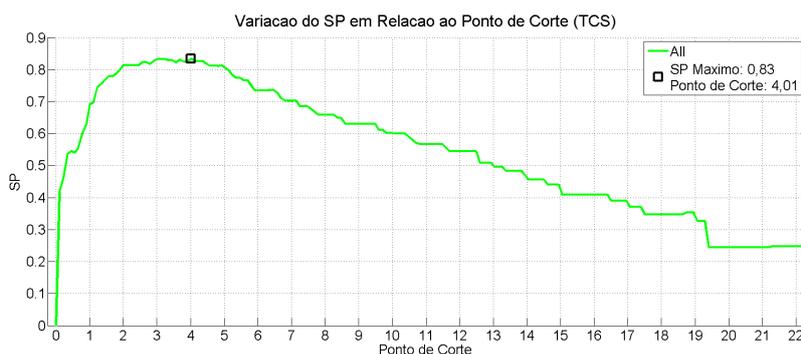
Tabela 5.32: Quantidades de Animais TB e Não TB na fazenda *All*

Animal	Quantidade
TB+	68
TB-	50
Anérgicos	7
<b>Total</b>	<b>125</b>

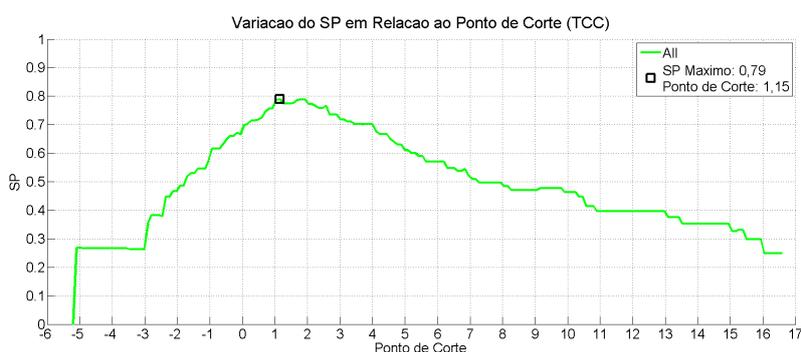
Os diferentes pontos de corte testados em cada exames foram considerados em faixas dinâmicas distintas. A faixa dinâmica encontrada para o TCS varia de 0 até 22,30 para  $\Delta\text{Bov}$ . Para o TCC, a faixa dinâmica dos valores de  $(\Delta\text{Bov} - \Delta\text{Av})$  varia de  $-5,20$  até  $16,70$ . Com tais intervalos definidos, cada teste cervical tem definidos variados pontos de corte e para cada um é testado e calculado o seu valor SP, gerando uma curva com a variação desta medida (Figura 5.39) ao longo da faixa dinâmica definida. Pela curva do TCS viu-se um aumento inicial vertiginoso do valor de SP, atingindo em seguida um valor máximo igual a  $0,89$  e nos cortes seguintes apresentou uma pequena estagnação seguido de queda no SP, atingindo um valor mínimo próximo de  $0,3$ . O ponto de corte associado à medida SP na curva é o corte de valor **4,01**, o mesmo encontrado na fazenda RS. A curva do TCC possui um comportamento semelhante, apresentando um aumento nos primeiros valores de corte testados e nos cortes seguintes não apresentando alterações. Próximo ao ponto de corte igual a  $-3$ , com a curva do TCC voltando apresentar crescimento até o seu pico de valor  $0,79$  de SP. Este valor máximo de SP foi obtido para o ponto de corte de valor **1,15** e após tal corte a curva apresenta uma queda gradual do valor SP, com uma queda abrupta ao final da faixa dinâmica. A Tabela 5.33 relaciona cada exame cervical com sua faixa dinâmica de valores, o ponto de corte ótimo e o valor SP associado a este corte.

Tabela 5.33: Valores de Faixa dinâmica, Pontos de Corte e SP para TCS e TCC da fazenda *All*

Exame	Faixa Dinâmica	Ponto de Corte	SP
TCS	[0 ; 22,30]	4,01	0,83
TCC	[-5,20 ; 16,70]	1,15	0,79



(a) Curva para o exame TCS



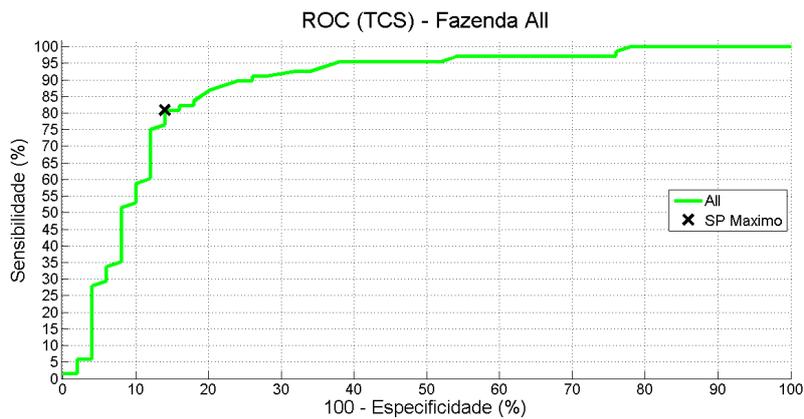
(b) Curva para o exame TCC

Figura 5.39: Variação do SP pelo ponto de corte para todas as Fazendas (*All*) para os exames TCS e TCC

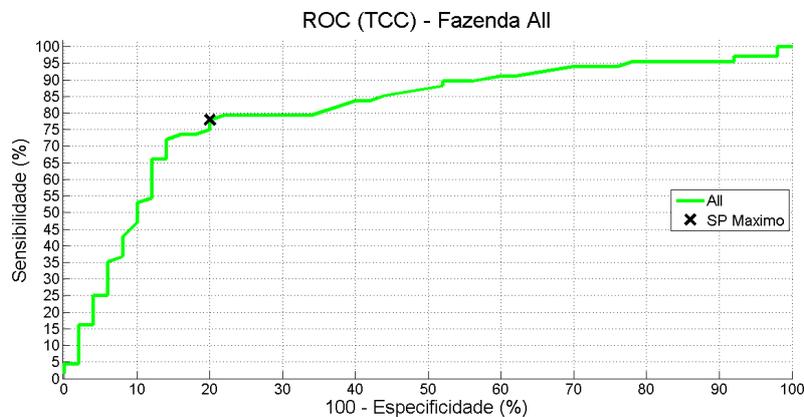
Comparada aos dados da outras fazendas, a Figura 5.39 não mostra aumentos ou quedas abruptas na variação do valor SP, indicando que não houve grandes variações de Sensibilidade ou Especificidade para novo ponto de corte testado. Tal característica indica que os dados da fazenda *All*, apesar de possuir uma faixa dinâmica mais extensa que as fazendas anteriores, possui esta preenchida com uma maior densidade e distribuição de animais, ou seja, as lacunas encontradas em cada fazenda foram preenchidas com a união delas.

Com os cortes definidos, o passo seguinte é a determinação de medidas de performance da fazenda. Através de análise de curva ROC (Figura 5.40) foi possível obter

as medidas de Sensibilidade e Especificidade dos testes. A Figura 5.40(a) exibe a curva ROC do TCS e um ponto destacado na curva, representando o ponto de corte de maior SP, indicando que tal ponto apresenta uma Sensibilidade de 81% e uma Especificidade de 86% na fazenda *All*. Para o TCC, a Figura 5.40(b) indica que no ponto de corte ótimo a eficiência da Sensibilidade é de 78% e na Especificidade é de 80%. Com estas medidas de performance e as quantidades de animais TB+ e TB- de cada uma das três fazendas, foram calculadas as Acuidades dos pontos de corte em cada teste cervical. O TCS apresentou uma Acuidade de 83%, enquanto o TCC obteve 79% de eficiência.



(a) Curva ROC do exame TCS



(b) Curva ROC do exame TCC

Figura 5.40: Curvas ROC da Fazenda All para os exames TCS e TCC

A Tabela 5.34 relaciona as medidas de performance de Sensibilidade, Especificidade, Acuidade e medida SP para cada um dos pontos de corte considerados nos testes cervicais. Com esta relação comparou-se a performance dos pontos de corte baseados no valor SP e dos cortes preconizados pela norma brasileira. A tabela trás a comparação apenas para os animais não anérgicos. Para a norma oficial, suas medidas de avaliação foram calculadas a partir dos dados das três fazendas.

Considerando o TCS, foi observado um ganho na Acuidade com o uso do novo ponto de corte, aumentando tal medida de 76% para 83% de eficiência, e tendo o valor SP também seguindo o mesmo comportamento, aumentando seu valor de 0,75 para a norma oficial até 0,83 com o corte baseado em SP. A Sensibilidade não apresentou alteração em sua eficiência e seu valor se manteve em 81%, enquanto a Especificidade apresentou um aumento de 16%, partindo de 70% até 86%. Já no TCC, novamente houve ganho nas eficiências da Acuidade, aumentando de 66% para 79%, e para o seu valor SP, que variou de 0,68 até 0,79. O teste apresentou um ganho de quase 30% em sua Sensibilidade, partindo de 51% para 78% de eficiência, porém com sua Especificidade apresentando uma perda de 6%, caindo de 86% para 80%.

Tanto a Acuidade quanto a medida SP são medidas dependentes da Sensibilidade e da Especificidade, explicando as variações vistas para os dois testes cervicais. No caso do TCS, o aumento de sua Acuidade foi causado somente pela Especificidade, já que a Sensibilidade não sofreu alteração. Para o TCC, houve mudanças tanto na Sensibilidade quanto na Especificidade, sendo que esta segunda medida apresentou uma pequena queda, o que em um primeiro momento impactou em perda na Acuidade, mas foi compensada em seguida pelo grande aumento na Sensibilidade.

Tabela 5.34: Eficiência das normas para exames cervicais para a fazenda All - animais não anérgicos

<b>Exame</b>	<b>Norma</b>	<b>Sens. (%)</b>	<b>Espec. (%)</b>	<b>Acc. (%)</b>	<b>SP</b>
TCS	PNCEBT	81	70	76	0,75
TCS	SP	81	86	83	0,83
TCC	PNCEBT	51	86	66	0,68
TCC	SP	78	80	79	0,79

Como complemento à análise da Tabela 5.34, a Figura 5.41 trás a distribuição dos animais não anérgicos e o destaque para cada ponto de corte considerado. Ao analisar a distribuição dos animais no TCS, o que se viu foi uma grande concentração de animais TB– na região de diagnóstico negativo da norma brasileira, seguida por uma pequena concentração destes animais na região de “Inconclusivo” e os demais animais dispersos na região de diagnóstico “Positivo”. A Sensibilidade do teste não foi alterada pois o novo corte e os cortes oficiais conseguem classificar os mesmos animais TB+, mas a Especificidade de ambos são diferentes pelo fato dos cortes oficiais não conseguirem classificar corretamente os animais TB– presentes na zona de “Inconclusivos”. Na análise do TCC, novamente se observa uma concentração de animais TB– na região de diagnóstico “Negativo”, enquanto nas regiões de “In-

conclusivo” e “Positivo” há uma maior concentração de bovinos TB+, sendo que em todos os casos há animais tanto doentes quanto sadios nas regiões citadas. O novo corte foi escolhido baseado no maior valor SP encontrado e foi definido em **1,15**. Este corte consegue identificar corretamente mais animais do que a norma oficial, como os animais TB+ concentrados na região de “Inconclusivos”, como visto pelo aumento de Acuidade e de Sensibilidade pelo corte. Logo, o que se viu pela Figura 5.41 foi que os ganhos obtidos pelos cortes baseado na medida SP na fazenda *All* se deu pela capacidade destes novos cortes conseguirem classificar corretamente animais presentes na região de diagnóstico “Inconclusivo” dos cortes preconizados pela norma brasileira.

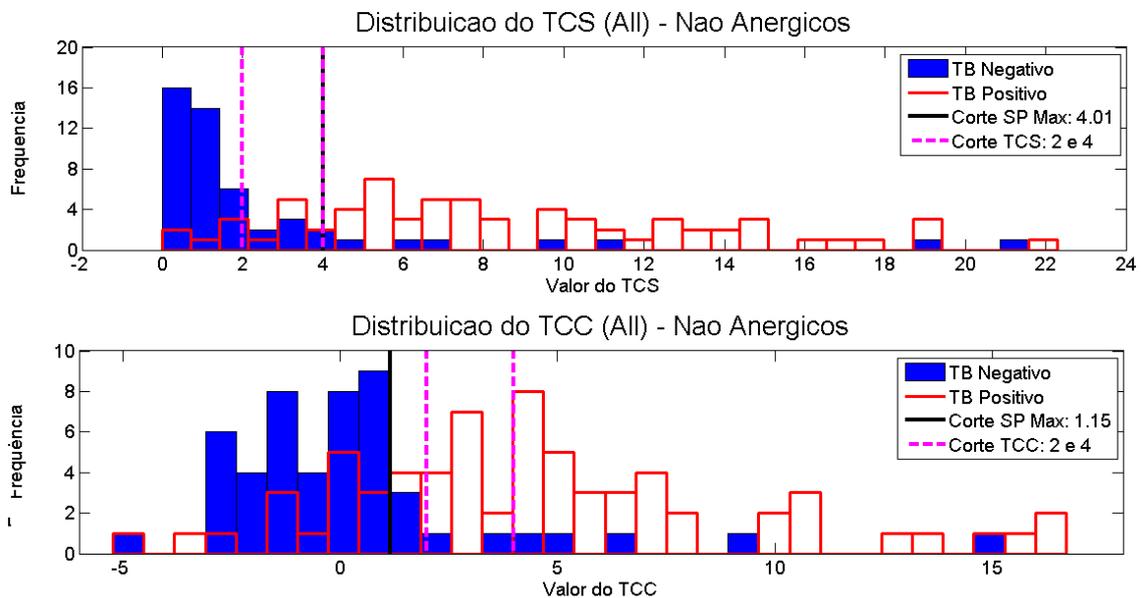


Figura 5.41: Testes Cervicais ma Fazenda *All* para animais não anérgicos

Para os animais anérgicos de *All*, a Tabela 5.35 relaciona a Sensibilidade obtida para estes animais por cada teste cervical e norma dos pontos de corte. O TCS apresentou uma eficiência de 14% (1 animal) em ambos os cortes considerados, enquanto que o TCC apresentou para a norma oficial do PNCEBT 0% de Sensibilidade e 57% (4 animais) com o ponto de corte otimizado.

Tabela 5.35: Eficiência das normas para exames cervicais para a fazenda All - animais anérgicos

Exame	Norma	Sensibilidade (%)
TCS	PNCEBT	14
TCS	SP	14
TCC	PNCEBT	0
TCC	SP	57

A Figura 5.42 mostra a distribuição dos animais anérgicos, para o TCS e TCC, e os cortes oficiais e otimizado por SP. Em TCS, 6 dos 7 animais anérgicos da fazenda estão presentes nas regiões de “Negativo” e “Inconclusivo” e uma vez que as normas consideradas possuem pontos de corte iguais a **4** e **4,01**, ambas conseguem apenas classificar o último animal mais à direita no gráfico. Tal animal classificado, é um dos anérgicos presentes originalmente na fazenda UFF, como visto na Seção 5.3.1. Os resultados do TCC para os animais anérgicos estão todos compreendidos nas regiões de diagnóstico “Negativo” e “Inconclusivo” da norma brasileira, explicando a razão para tal norma possuir 0% como taxa de Sensibilidade.

Já o ponto de corte otimizado é capaz de identificar 4 animais, uma vez que seu valor é baixo (**1,15**). Já que, quanto menor o ponto de corte maior a Sensibilidade do teste cervical, a chance de identificar um animal com anergia aumenta.

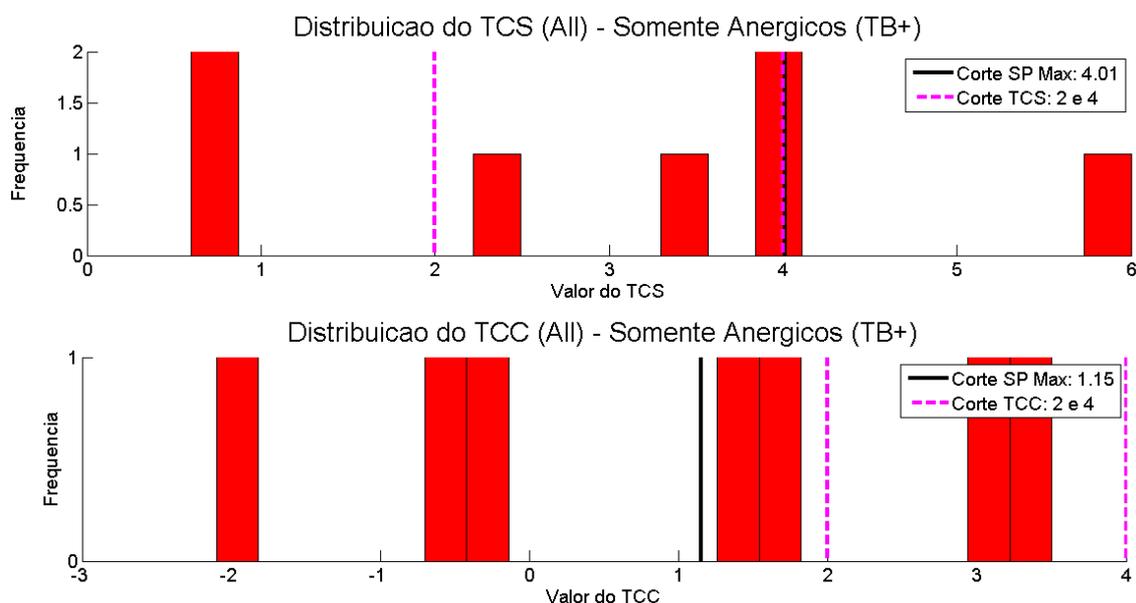


Figura 5.42: Testes Cervicais na Fazenda All para animais anérgicos

Pelos resultados expostos nesta seção o que se viu foi, que um ponto de corte baseado em SP apresentou ganhos na eficiência de ambos os testes cervicais. Nos

animais não anérgicos foi observado aumento na eficiência tanto no TCS quanto no TCC. Nos animais anérgicos, ainda que estes bovinos não reajam positivamente à estes testes, foi observado que é possível identificar uma parcela destes animais com a norma proposta, como visto somente no TCC. A abordagem global adotada neste experimento, mostrou que a normal oficial utilizada possui carências na Especificidade do TCS e na Sensibilidade do TCC.

# Capítulo 6

## Conclusões

Neste trabalho foi abordado o problema que a Tuberculose Bovina representa hoje no Brasil, que além de ser considerado um problema econômico pode também ser visto como problema de saúde pública. A erradicação da doença no Brasil é o principal objetivo do PNCEBT, programa que determina as regras para combatê-la, mas que não vem sendo capaz de identificar totalmente os animais contaminados e prevenir a disseminação da doença.

A proposta deste estudo foi avaliar novas possibilidades que possam gerar aumentos na eficiência do programa, e assim foi considerado o uso de exames complementares ao já preconizados pelo PNCEBT, como também um estudo sobre os testes cervicais hoje usados.

Este capítulo apresenta as conclusões obtidas com os experimentos conduzidos neste estudo e propostas de trabalhos futuros.

### 6.1 Adição de exames e Repetibilidade

A primeira proposta neste estudo foi avaliar o uso de uma abordagem multi-diagnóstica para a identificação da Tuberculose Bovina no Brasil. A norma hoje vigente determina que apenas um tipo de exame, os testes cervicais, seja utilizado na identificação da doença porém, estes exames apresentam limitações e não conseguem identificar todos os bovinos tuberculosos presentes nos rebanhos do país. A abordagem multi-diagnóstica faz uso de um conjunto maior de exames, que venham a identificar com mais eficiência a Tuberculose nos bovinos.

A partir de conjuntos de dados de diferentes fazendas, onde para cada fazenda havia uma variedade de exames complementares aos testes cervicais, foi feito um estudo sobre o uso destes novos exames na identificação da Tuberculose, em conjunto com os testes oficiais, e foi analisado se o acréscimo desses exames era benéfico ao diagnóstico de TB. Tal estudo foi feito por uma abordagem neural, onde foram

desenvolvidos classificadores neurais com a função de determinar se um animal era ou não tuberculoso, de acordo com os resultados de seus exames.

Foram utilizados dados referentes à três fazendas distintas (UFF, Macaé e RS), e nelas foram encontrados dois exames complementares. O primeiro deles foi o ELISA, um exame sorológico que identifica a presença de anticorpos específicos para a Tuberculose a partir de amostras de soro do bovino. O segundo exame considerado busca identificar *M. bovis*, o agente causador da doença, em amostras de leite retirada dos animais.

O estudo feito sobre o exame de leite apresentou limitações, pois este está presente apenas nos dados referentes à fazenda UFF e nem todos os bovinos da fazenda tiveram o exame realizado. No entanto, mesmo com essa restrição, o exame se mostrou promissor e com bons resultados. Os modelos neurais desenvolvidos, ao terem o exame de leite adicionados aos testes cervicais, apresentaram um aumento de performance nas medidas de Sensibilidade e Especificidade. Esse ganho no entanto foi maior na medida de Sensibilidade, sendo este um indicador que o exame é mais apropriado para se identificar animais TB+. Ao considerar o exame de leite somente para animais anérgicos, nada pode-se dizer sobre o seu uso, já que em nenhum dos animais anérgicos da fazenda ele foi aplicado. Logo, pelos resultados vistos na fazenda UFF, a adição do exame de leite aos testes cervicais apresentou impactos positivos na identificação da doença.

Já para ELISA o seu estudo foi mais abrangente, uma vez que este tipo de exame esteve presente em todas as fazendas, ainda que com variedades diferentes do mesmo. O teste ELISA apresentou impactos na identificação de animais não anérgicos nos dados de UFF e Macaé, enquanto que em RS não foram observadas mudanças com a adição do exame. ELISA é um exame com alta Especificidade em bovinos não anérgicos e o seu uso em conjunto com os testes cervicais apresentou ganho para esta medida em ambas as fazendas. Para a Sensibilidade, foi observado um pequeno ganho apenas na fazenda Macaé, indicando que os ganhos com o exame não se limitam à Especificidade.

ELISA possui ainda a aplicação específica de identificação de animais anérgicos. Tal característica foi observada nas fazendas UFF e Macaé, as únicas que apresentaram este tipo de animal, e foram observados ganhos na Sensibilidade dos modelos para este tipo de bovino.

Considerando ainda o uso conjunto dos testes cervicais e ELISA apenas, foi realizado um experimento de Repetibilidade e nele os dados de RS foram testados nos modelos neurais desenvolvidos com dados de UFF e Macaé. A fazenda RS apresenta limitações em seus dados, como o preenchimento incompleto de exames, uma quantidade pequena de amostras de bovinos, tal como uma grande diferença nas ocorrências de animais TB+ e TB-, que acabaram influenciado os resultados tanto

no teste de seus modelos neurais como na repetibilidade. No entanto, o experimento apontou que o comportamento dos animais TB+ de RS esteve dentro da variação estatística dos modelos de UFF. Tal fato indica que os mesmos impactos vistos nos animais TB+ de UFF se repetem para nestes animais de RS.

A fazenda UFF foi a única em que se pode testar o uso conjunto do exame de leite, de ELISA e dos testes cervicais. Para os modelos testados, o uso conjunto dos quatro exames apresentou as melhores medidas de performance, acarretando em aumento de Sensibilidade e Especificidade para animais não anérgicos. Para os animais anérgicos, o uso conjunto dos exames apresentou ganhos novamente na Sensibilidade, do que se comparado aos testes cervicais apenas. Logo, o uso combinado dos dois exames também acarretaram em ganhos na identificação da Tuberculose.

Os experimentos conduzidos neste estudo mostraram que o uso da abordagem multi-diagnóstica aumenta a eficiência da identificação de Tuberculose bovina. Tanto o exame de leite quanto ELISA são exames candidatos a serem usados em conjunto com os testes cervicais e a conclusão deste estudo é que ambos devem ser usados em conjunto com os testes cervicais para se obter a melhor performance na identificação de Tuberculose. No entanto, cada exame apresenta fatores de custo para sua adição pelo PNCEBT, como por exemplo: material usado, mão de obra para aplicação do exame, logística para avaliação do resultado. Entendendo que não necessariamente a adição de ambos os exames é possível, este estudo propõe a seguinte ordem para adição dos exames:

1. ELISA
2. Leite

O leite, apesar dos resultados promissores, foi preterido como segundo exame a ser adicionado por ainda precisar de outros estudos que confirmem seus resultados, tal como o seu uso para animais anérgicos. Já ELISA foi escolhido como primeiro exame a ser adicionado por apresentar ganhos na identificação de TB mas principalmente por ser o exame indicado para a identificação de animais anérgicos. Uma vez que animais anérgicos são animais que não reagem positivamente aos exames hoje empregados na norma brasileira, ainda que estejam doentes, tais animais se tornam fonte transmissora da doença e podem condenar todo um rebanho ao abate.

## 6.2 Testes Cervicais

Neste trabalho também foi estudado e discutido separadamente apenas os testes cervicais, com o objetivo de avaliar as normas oficiais, que definem o seu uso e interpretação, além de propor uma nova norma, de modo a melhorar a eficiência dos

mesmos. A proposta para os testes cervicais tinha como ponto principal a definição de novos pontos de cortes ou limites que determinassem se um animal é positivo ou negativo para Tuberculose Bovina. A nova norma proposta primeiramente desconsiderava o resultado “Inconclusivo” e conseqüentemente apresentou apenas um único ponto de corte, que era escolhido sobre o valor SP. A análise foi conduzida sobre as três fazendas separadamente e por fim sobre todas as fazendas juntas, em uma abordagem global e que pudesse se aproximar do cenário da Tuberculose Bovina no Brasil. Para as fazendas de UFF, Macaé, RS e a fazenda fictícia *All* (união das três fazendas) o uso do novo ponto de corte baseado no valor SP apresentou aumento na eficiência de ambos os testes cervicais, melhorando a Sensibilidade, a Especificidade ou ainda ambas medidas dos testes cervicais, dependendo da fazenda estudada.

Foi concluído que o maior ponto fraco dos cortes preconizados pela normal oficial é o seu diagnóstico “Inconclusivo” para os bovinos, e que com a simples remoção desse diagnóstico é possível obter ganhos na eficiência dos testes. Este diagnóstico não representa um dos possíveis estados de um animal para a Tuberculose, onde este pode apenas estar ou não estar infectado pelo agente *M. bovis*. Os resultados encontrados em cada fazenda mostraram que, dependendo do teste cervical analisado, pode haver mais animais doentes ou sadios que são classificados erradamente como inconclusivos.

Sobre cada exame cervical, observou-se que o exame TCC sistematicamente apresentou, em todos os casos estudados, um ponto de corte inferior ao valor hoje usado, sempre gerando um aumento na Sensibilidade. Esta característica sugere que o TCC é um exame pouco sensível na identificação de animais positivos para Tuberculose e que uma mudança em sua norma deve buscar aumentar a Sensibilidade do exame. Para a análise global, encontrou-se que seu ponto de corte foi **1,15** e isso indica que a norma oficial vigente usa um corte demasiadamente alto e com baixa Sensibilidade. A possibilidade de ajuste da norma vigente passaria diretamente pela alteração de seu corte para um valor mais baixo.

Para o exame TCS, foi observado que os cortes variaram de acordo com a fazenda, podendo apresentar valores menores, maiores ou mesmo iguais aos cortes oficiais. Na abordagem global deste teste cervical, se averiguou que os cortes oficiais possuem baixa Especificidade e que um corte igual **4,01** aumentaria tal medida e manteria a Sensibilidade. O ganho observado para tal exame foi conseguido pela exclusão do diagnóstico inconclusivo. Os resultados encontrados sugerem que o TCS é um exame cuja sua deficiência é a classificação dos animais negativos para Tuberculose.

Apesar dos animais anérgicos não serem facilmente detectáveis pelos testes cervicais, a eliminação do diagnóstico “Inconclusivo” e um ponto de corte menor é capaz de aumentar a identificação desses animais pelos exames, como visto para o TCC.

Logo, a conclusão deste estudo é que os testes cervicais podem ser melhorados,

apresentando uma maior eficiência na separação de animais, com ou sem Tuberculose, e que isto pode ser alcançado através de uma revisão das normais oficiais que hoje são estabelecidas.

### 6.3 Trabalhos Futuros

A abordagem multi-diagnóstica apresentou bons resultados e significativo aumento na taxa de diagnóstico de Tuberculose Bovina. Contudo, o estudo conduzido contou somente com ELISA e o teste do leite como exames complementares, sendo que este segundo exame estava presente em apenas uma fazenda e em uma quantidade reduzida dos dados. Para trabalhos futuros, sugerimos repetir a análise deste trabalho para a adição do exame de leite aos testes cervicais, porém com um conjunto de dados com mais ocorrências de animais com este exame feito. Nesta nova análise, buscar avaliar a performance do exame também para os animais anérgicos.

Uma segunda possibilidade para trabalhos futuros é aumentar a variedade de exames adicionados aos testes cervicais. Logo, a proposta é repetir o experimento conduzido neste estudo com outros exames complementares, como por exemplo o exame para o muco nasal do bovino ou o uso de novos ELISAs recombinantes.

Neste trabalho a adição dos exames junto aos testes cervicais foi avaliada baseando-se apenas nos impactos causados na eficiência do diagnóstico da Tuberculose. Sendo assim, um estudo que pode ser feito é o de custo efetivo para a implantação de cada exame junto a norma a brasileira, com o objetivo verificar se sua aplicação é viável economicamente.

Em uma análise focada apenas nos testes cervicais, uma proposta é utilizar uma maior variedade de fazendas e uma maior quantidade de animais, já avaliados por seus testes cervicais, para repetir a abordagem global para um novo ponto de corte para os testes intradérmicos. Ao utilizar uma maior variedade de animais e fazendas, mais próximo é este cenário do problema da Tuberculose Bovina no Brasil e mais confiável serão os novos cortes propostos.

Uma última proposta para trabalhos futuros é o desenvolvimento de novos classificadores neurais a serem empregados ao problema da Tuberculose Bovina. Os resultados promissores que os Perceptrons Multi Camadas apresentaram neste estudo gera a possibilidade de aplicação desses modelos em campo e em cenários práticos. Esses novos classificadores podem ser usados como apoio ao diagnóstico da doença.

# Referências Bibliográficas

- [1] LEVADA, M. M. O., FIERI, W. J., PIVESSO, M. S. G. *Apontamentos Teóricos de Citologia, Histologia e Embriologia*. 5 ed. São Paulo, Catálise Editora, 1996.
- [2] FILHO, E. S. *Análise Não-Linear de Componentes Independentes para uma Filtragem Online Baseada em Calorimetria de Alta Energia e com Fina Segmentação*. Tese de D.Sc., COPPE/UFRJ, Rio de Janeiro, 2010.
- [3] AZEVEDO, S. *Instrução Normativa 2/2001. Programa Nacional de Controle e Erradicação da Brucelose e da Tuberculose*. 2 ed. Brasília, EMBRAPA, 2001.
- [4] PRITCHARD., D. G. “A Century of Bovine Tuberculosis 1888-1988: Conquest and Controversy”, *Journal of Comparative Pathology*, v. 99, n. 4, pp. 357–399, 1998.
- [5] CRAWSHAW, T. “Wildlife Hosts for *Mycobacterium bovis*”, *Veterinary Record*, v. 67, n. 11, pp. 164–165, 2013.
- [6] HAYKIN, S. *Redes Neurais: Princípio e Prática*. 2 ed. Porto Alegre, Bookman, 2008.
- [7] WASSERMAN, P. *Neuro Computing: Theory and Practice*. 1 ed. Nova Iorque, Coriolis Group, 1989.
- [8] HAVEL, J., AMATO, F., LOPEZ, A., et al. “Artificial Neural Networks in Medical Diagnosis”, *Journal of Applied Biomedicine*, v. 11, n. 2, pp. 47–58, 2013.
- [9] SEIXAS, J. M., FARIA, J., VIEIRA, A. F., et al. “Artificial Neural Network Models to Support The Diagnosis of Pleural Tuberculosis in Adult Patients”, *The International Journal of Tuberculosis and Lung Disease*, v. 17, n. 5, pp. 682–686, 2012.

- [10] DE OLIVEIRA COSTA, F., MOTTA, L. C. S., NOGUEIRA, J. L. T. “Uma abordagem baseada em Redes Neurais Artificiais para o auxílio ao diagnóstico de doenças meningocócicas”, *Revista Brasileira de Computação Aplicada*, v. 2, n. 1, pp. 79–88, 2010.
- [11] SIRISTATIDIS, C., CHRELIAS, C., POULIAKIS, A., et al. “Artificial neural networks in gynaecological diseases: current and potential future applications”, *Medical Science Monitor*, v. 16, n. 10, pp. 231–236, 2010.
- [12] ROWLEY, H. A., BALUJA, S., KANADE, T. “Neural Network-Based Face Detection”, *Transactions On Pattern Analysis And Machine Intelligence*, v. 20, n. 1, pp. 23–38, 1998.
- [13] CIODARO, T., DEVA, D., SEIXAS, J. M., et al. “Online Particle Detection With Neural Networks Based on Topological Calorimetry Information”, *Journal of Physics*, v. 20, n. 1, pp. 23–38, 2012.
- [14] PATIDAR, R., SHARMA, L. “Credit Card Fraud Detection Using Neural Network”, *International Journal of Soft Computing and Engineering*, v. 1, n. 1, pp. 32–38, 2011.
- [15] MARASSI, C. D., MEDEIROS, L., LILENBAUM, W., et al. “A Multidisciplinary Approach to Diagnose Naturally Occurring Bovine Tuberculosis in Brazil”, *Pesquisa Veterinária Brasileira*, v. 33, n. 1, pp. 15–20, 2013.
- [16] DOMENECH, R. R., GOODCHILD, A. T., VORDERMEIER, H. M., et al. “*Ante Mortem* Diagnosis of Tuberculosis in Cattle: A Review of The Tuberculin Tests, Gamma-Interferon Assay and Other Ancillary Diagnostic Techniques”, *Research in Veterinary Science*, v. 81, n. 2, pp. 190–210, 2006.
- [17] JORBAN, M. I., MITCHELL, T. M. “Machine Learning: Trends, Perspectives, and Prospects”, *Science*, v. 349, n. 6245, pp. 255–260, 2015.
- [18] ABRAHAO, R. M. C. “Tuberculose Humana Causada pelo *Mycobacterium bovis*: Considerações Gerais e a Importância dos Reservatórios Animais”, *Archives of Veterinary Science*, v. 4, n. 1, pp. 5–15, 1999.
- [19] CASTRO, K. *Tuberculose Bovina*. Dissertação de M.Sc., Universidade Castelo Branco, Vitória, 2008.
- [20] ALMEIDA, R. F. C., SOARES, C. O., ARAUJO, F. R. *Brucelose e Tuberculose Bovina: Epidemiologia, Controle e Diagnóstico*. 1 ed. Brasília, EMBRAPA, 2004.

- [21] CORNER, L., MELVILLE, L., MCCUBBIN, K., et al. “Efficiency of Inspection Procedures for The Detection of Tuberculous Lesions in Cattle”, *Australian Veterinary Journal*, v. 67, n. 11, pp. 389–392, 1990.
- [22] CORREA, W. M., CORREA, C. N. M. *Enfermidades Infecciosas dos Mamíferos Domésticos*. 2 ed. Rio de Janeiro, Medsi, 1992.
- [23] RADOSTITS, O. M., GAY, C. C., BLOOD, D. C., et al. *Clínica Veterinária: Um Tratado de Doenças dos Bovinos, Ovinos, Suínos, Caprinos e Equinos*. 9 ed. Brasil, Guanabara Koogan, 2002.
- [24] WHIPPLE, D. L., BOLIN, C. A., MILLER, J. M. “Distribution of Lesions in Cattle Infected With *Mycobacterium bovis*”, *Journal of Veterinary Diagnostic Investigation*, v. 8, n. 3, pp. 351–354, 1996.
- [25] OLIVEIRA, L. B., BARRETO, I. D. C., VIEIRA, S. D., et al. “Prospecção Tecnológica Sobre o Controle e Combate da Tuberculose Bovina”, *International Symposium on Technological Innovation*, v. 2, n. 1, pp. 262–268, 2014.
- [26] POLLOCK, J. M., NEILL, S. D. “*Mycobacterium bovis* Infection and Tuberculosis in Cattle”, *Veterinary Journal*, v. 1, n. 163, pp. 115–127, 2002.
- [27] MENDEZ, M. C., LEMOS, R. A. A., SCHILD, A. L., et al. *Doenças de Ruminantes e Equinos*. 2 ed. Brasil, Varela, 2001.
- [28] LENQUIN, R. M. “Enzyme Immunoassay (EIA)/Enzyme-Linked Immunosorbent Assay (ELISA)”, *Clinical Chemistry*, v. 51, n. 12, pp. 2415–2418, 2005.
- [29] SAAH, A. J., FARZADEGAN, H., FOX, R., et al. “Detection of Early Antibodies in Human Immunodeficiency Virus Infection by Enzyme-Linked Immunosorbent Assay, Western Blot, and Radioimmunoprecipitation”, *Journal Of Clinical Microbiology*, v. 25, n. 9, pp. 1605–1610, 1987.
- [30] SUPRANB, E. M., OXALL, E. H., CRASKE, J., et al. “Enzyme-Linked Immunosorbent Assay (ELISA) for The Detection of Hepatitis Be Antigen and Antibody: Report of a Field Trial”, *Journal of Clinical Pathology*, v. 36, n. 5, pp. 581–585, 1983.
- [31] SILVA, M. R., PELLEGRIN, A. O., ARAUJO, F. R., et al. “Avaliação de Proteínas Recombinantes para Detecção de Anticorpos Contra *Mycobacterium bovis* por ELISA”, *Boletim de Pesquisa e Desenvolvimento*, v. 1, n. 34, 2015.

- [32] EHLICH, P. “A Method for Staining The Tubercle Bacillus”, *Deutsche Medizinische Wochenschrift*, v. 8, n. 1, pp. 269–270, 1882.
- [33] EZEMBRO, E., AZAM, K., CADIR, N. *Manual de Baciloscopia da Tuberculose*. 1 ed. Moçambique, Programa Nacional de Controle da Tuberculose, 2012.
- [34] TEIXEIRA, H. C., ABRAMO, C., MUNK, M. E. “Diagnóstico Imunológico da Tuberculose: Problemas e Estratégias para o Sucesso”, *Jornal Brasileiro de Pneumologia*, v. 33, n. 3, pp. 323–334, 2007.
- [35] FRÁGUAS, S. A., ABREU, M. S. C., LILENBAUN, W., et al. “Estudo Comparativo de Métodos Complementares para o Diagnóstico da Tuberculose Bovina em Animais Reagentes à Tuberculinização”, *Revista Brasileira Ciência Veterinária*, v. 15, n. 3, pp. 117–121, 2008.
- [36] ROXO, E. “Tuberculose Bovina: Revisão”, *Arquivo do Instituto Biológico*, v. 63, n. 2, pp. 91–97, 1996.
- [37] SIMPSON, D., REED, R. “Slide Culture of Tubercle Bacilli. IV: Further Diagnostic Trials”, *American Journal of Public Health and Nation’s Health*, v. 48, n. 9, pp. 1153–1161, 1958.
- [38] KRITSKI, A. L., CONDE, M. B., SOUZA, G. R. M. *Tuberculose: do Ambulatório à Enfermaria*. 3 ed. Rio de Janeiro, Editora Atheneu Rio, 2000.
- [39] KONEMAN, E., ALLEN, S., JANDA, W., et al. *Diagnóstico Microbiológico*. 6 ed. Rio de Janeiro, Guanabara Koogan, 2001.
- [40] FUVERKI, R. B. N., MURAKAMI, P. S., BIONDO, A. W., et al. “Uso da PCR para Detecção e Identificação de Microbactérias a Partir de Amostras Clínicas de Bovinos”, *Archives of Veterinary Science*, v. 13, n. 1, pp. 73–77, 2008.
- [41] WARDS, B. J., COLLINS, D. M., LISLE, G. W. “Detection of *Mycobacterium bovis* in Tissues by Polimerase Chain Reation”, *Veterinary Microbiology*, v. 43, pp. 227–240, 1995.
- [42] ZANINI, M. S., MOREIRA, E. C., LOPES, M. T. P., et al. “*Mycobacterium bovis* Polymerase Chain Reaction Identification in Bovine Lynphnode Biopsies and Genotyping in Isolates From Southeast Brazil by Apolygotyping and Restriction Fragment Lenght Polymorphism”, *Memórias Instituto Oswaldo Cruz*, v. 96, n. 6, 2001.

- [43] RUGGIERO, A. P., IKUNO, A. A., FERREIRA, V. C. A., et al. “Tuberculose Bovina: Alternativas Para o Diagnóstico”, *Arquivo Instituto Biológico*, v. 74, n. 1, pp. 55–65, 2007.
- [44] KLATSER, P. R., KUIJPER, S., VAN, C. W. I., et al. “Stabilized, Freeze-Dried PCR Mix for Detection of *Mycobacteria*”, *Journal of Clinical Microbiology*, v. 36, n. 6, pp. 798–800, 1998.
- [45] ARAUJO, C. P., OSORIO, A. L. A. R., JORGE, K. S. G., et al. “Direct Detection of *Mycobacterium tuberculosis* Complex in Bovine and Bubaline Tissues Through Nested-PCR”, *Sociedade Brasileira de Microbiologia*, v. 45, n. 2, pp. 633–640, 2014.
- [46] ARAUJO, C. P., OSORIO, A. L. A. R., ARAUJO, F. R., et al. “Detection of *Mycobacterium bovis* in Bovine and Bubaline Tissues Using Nested-PCR for TbD1”, *PLOS ONE*, v. 9, n. 3, pp. 1–6, 2014.
- [47] MEDEIROS, L. *Abordagem Multidisciplinar No Diagnóstico da Tuberculose Bovina*. Dissertação de M.Sc., Universidade Federal Fluminense, Rio de Janeiro, 2009.
- [48] SOUZA, M. *Tuberculose Bovina: Diagnóstico Intradérmico e Exames Complementares em Propriedade de Exploração Leiteira*. Dissertação de M.Sc., Universidade Federal de Uberlândia, Uberlândia, 2013.
- [49] HOULIHAN, M. G., DIXON, F. W., PAGE, N. A. “Outbreak of Bovine Tuberculosis Featuring Anergy to the Skin Test, Udder Lesions and Milkborne Disease in Young Calves”, *Veterinary Record*, v. 163, n. 12, pp. 357–361, 2008.
- [50] PLACKETT, P., RIPPER, J., CORNER, L., et al. “An ELISA for The Detection of Anergic Tuberculous Cattle”, *Australian Veterinary Journal*, v. 66, n. 1, pp. 15–19, 1989.
- [51] LILENBAUM, W., FONSECA, L. S. “The Use of ELISA as a Complementary Tool for Bovine Tuberculosis Control in Brazil”, *Brazilian Journal of Veterinary Research and Animal Science*, v. 43, n. 2, pp. 256–261, 2006.
- [52] FURNALETTO, L. V., FIGUEIREDO, E. E. S., LILENBAUM, W., et al. “Prevalência de Tuberculose Bovina em Animais e Rebanhos Abatidos em 2009 no Estado de Mato Grosso, Brasil”, *Arquivo Brasileiro de Medicina Veterinária e Zootecnia*, v. 64, n. 2, pp. 274–280, 2012.

- [53] SOUSA, C. L. M. “Considerações Sobre a Tuberculose Bovina no Norte Fluminense e no Município de Campos dos Goytacazes Após Advento do PNCEBT - Programa Nacional de Controle e Erradicação da Brucelose e Tuberculose Bovina”, *Perspectivas Online*, v. 2, n. 8, pp. 138–148, 2008.
- [54] FIGUEIREDO, S. M., ROCHA, V. C. M., HIGINO, S. S., et al. “Tuberculose Bovina no Estado da Paraíba: Estudo Retrospectivo”, *Pesquisa Veterinária Brasileira*, v. 30, n. 9, pp. 712–716, 2010.
- [55] SILVA, D. A. V., BURGER, K. P., MARTINS, A. M. C. V., et al. “Identificação de Lesões Macroscópicas Sugestivas de Tuberculose Bovina”, *Revista Brasileira de Higiene e Sanidade Animal*, v. 8, n. 2, pp. 149–160, 2014.
- [56] KANTOR, I. N., RITACCO, V. “Bovine Tuberculosis in Latin America and Caribbean: Current Status, Control and Eradication Programs”, *Veterinary Microbiology*, v. 40, n. 1, pp. 5–14, 1994.
- [57] JORGE, M. H. P. M., GOTLIEB, S. L. D., LAURENTI, R. *A Saúde no Brasil: Análise do Período 1996-1999*. 1 ed. Brasília, Editora Parma Ltda, 2001.
- [58] MARSLAND, S. *Machine Learning - An Algorithmic Perspective*. 1 ed. Florida, CRC Press, 2009.
- [59] BISHOP, C. M. *Pattern Recognition and Machine Learning*. 1 ed. Cambridge, Reino Unido, Springer, 2006.
- [60] WEBB, A. R. *Statistical Pattern Recognition*. 2 ed. Malvern, Reino Unido, John Wiley & Sons, 2002.
- [61] AGUIAR, F., PINTO, J., TORRES, R., et al. “Artificial Neural Network (ANN) for Prediction of Pulmonary Tuberculosis in Hospitalized Patients”, *European Respiratory Journal*, v. 42, n. 57, pp. 32–38, 2013.
- [62] MOSTAFA, Y., ISMAIL, M. M., LIN, H. T. *Learning From Data - A Short Course*. 1 ed. Estados Unidos da América, AMLBook, 2012.
- [63] CIBENKO, G. “Approximation by Superpositions of a Sigmoidal Function”, *Mathematics of Control, Signals and Systems*, v. 2, n. 4, pp. 303–314, 1989.
- [64] ZHANG, H., WANG, Z., LIU, D. “A Comprehensive Review of Stability Analysis of Continuous-Time Recurrent Neural Networks”, *IEEE Transactions on Neural Networks and Learning Systems*, v. 25, n. 7, pp. 1229–1262, 2014.

- [65] FAHLAND, D., AALST, W. M. P. V. D. “Simplifying Discovered Process Models in a Controlled Manner”, *Information Systems*, v. 38, n. 4, pp. 585–605, 2013.
- [66] DOMINGOS, P. “A Few Useful Things to Know About Machine Learning”, *Communications of the ACM*, v. 55, n. 10, pp. 78–81, 2012.
- [67] KARSOLIYA, S. “Approximating Number of Hidden Layer Neurons in Multiple Hidden Layer BPNN Architecture”, *International Journal of Engineering Trends and Technology*, v. 3, n. 6, pp. 714–717, 2012.
- [68] REIDMILLER, M. “Advanced Supervised Learning in Multi-layer Perceptrons - From Backpropagation to Adaptive Learning Algorithms”, *Computer Standards & Interfaces*, v. 16, n. 3, pp. 265–278, 1994.
- [69] RIEDMILLER, M., BRAUN, H. “A Direct Adaptive Method for Faster Backpropagation Learning: The RPROP Algorithm”, *Neural Networks*, v. 1, pp. 586–591, 1993.
- [70] RUMELHART, D. E., HINTON, G. E., WILLIAMS, R. J. “Learning representations by back-propagating errors”, *Nature*, v. 323, n. 6088, pp. 533–536, 1986.
- [71] POWERS, D. M. W. “Evaluation: From Precision, Recall and F-Measure to ROC, Informedness, Markedness & Correlation”, *Journal of Machine Learning Technologies*, v. 2, n. 1, pp. 37–63, 2011.
- [72] REISER, B., FARAGOCCI, D. “Confidence Intervals for The Generalized ROC Criterion”, *International Biometric Society*, v. 2, n. 53, pp. 644–652, 1997.
- [73] FAWCETT, T. “An introduction to ROC analysis”, *Pattern Recognition Letters*, v. 27, n. 8, pp. 861–874, 2006.
- [74] KUMAR, R., INDRAYAN, A. “Receiver Operating Characteristic (ROC) Curve for Medical Researchers”, *Indian Pediatrics*, v. 48, n. 4, pp. 277–287, 2011.
- [75] ZANGIACOMI, E., LOUZADA, F., BRAGANÇA, B. “A Curva ROC para Testes Diagnósticos”, *Caderno Saúde Coletiva*, v. 1, n. 11, pp. 7–31, 2003.
- [76] FLETCHER, R. W., FLETCHER, S. E., FLETCHER, G. S. *Epidemiologia Clínica*. 5 ed. Porto Alegre, Artmed, 2014.

- [77] ANJOS, A., TORRES, R. C., SEIXAS, J. M., et al. “Neural Triggering System Operating on High Resolution Calorimetry Information”, *Nuclear Instruments and Methods in Physics Research A*, v. 559, n. 1, pp. 134–138, 2006.
- [78] KOHAVI, R. “A Study of CrossValidation and Bootstrap for Accuracy Estimation and Model Selection”, *International Joint Conference on Artificial Intelligence*, v. 2, n. 14, pp. 1137–1143, 1995.
- [79] RAMANAKUMAR, A. V., THOMANN, P., CANDEIAS, J. M., et al. “Use of The Normalized Absorbance Ratio as an Internal Standardization Approach To Minimize Measurement Error in Enzyme-Linked Immunosorbent Assays for Diagnosis of Human Papillomavirus Infection”, *Journal Of Clinical Microbiology*, v. 48, n. 3, pp. 791–796, 2010.
- [80] WONG, R. C. W., FAVALORO, E. J., POLLOCK, W., et al. “A Multi-Centre Evaluation of The IntraAssay and InterAssay Variation of Commercial and InHouse AntiCardiolipin Antibody Assays”, *Pathology*, v. 36, n. 2, pp. 182–192, 2004.
- [81] FLORAS, A. N. K., HOLOWAYCHUK, M. K., HODGINS, D. C., et al. “Investigation of a Commercial ELISA for The Detection of Canine Procalcitonin”, *Journal of Veterinary Internal Medicine*, v. 28, n. 2, pp. 599–602, 2014.
- [82] LILENBAUM, W., RIBEIRO, E. R., SOUZA, G. N., et al. “Evaluation of an ELISA-PPD for The Diagnosis of Bovine Tuberculosis in Field Trials in Brazil”, *Research in Veterinary Science*, v. 66, n. 3, pp. 191–195, 1999.
- [83] HUANG, G., SONG, S., YOU, K. “Trends in Extreme Learning Machines: A Review”, *Neural Networks*, v. 61, pp. 32–48, 2015.
- [84] HORNIK, K., STINCHCOMBE, M., WHITE, H. “Multilayer Feedforward Networks Are Universal Approximators”, *Neural Networks*, v. 2, n. 2, pp. 359–366, 1989.
- [85] SHAO, Y. “Comparison of Early Stopping Criteria for Neural-Network-Based Subpixel Classification”, *Geoscience and Remote Sensing Letters*, v. 8, n. 1, pp. 113–117, 2011.
- [86] INTERNATIONAL, A. “ASTM E177-14 Standard Practice for Use of the Terms Precision and Bias in ASTM Test Methods1”, *Geoscience and Remote Sensing Letters*, v. 8, n. 1, pp. 113–117, 2014.

- [87] RAMOS, D. F., SILVA, A. B. S., FAGUNDES, M. Q., et al. “Molecular Typing of *Mycobacterium bovis* Isolated in the South of Brazil”, *Brazilian Journal of Microbiology*, v. 45, n. 2, pp. 657–660, 2014.
- [88] PERUMAALLA, V. S., ADAMS, L. G., PAYEUR, J., et al. “Molecular Fingerprinting Confirms Extensive Cow-to-cow Intra-herd Transmission of a Single *Mycobacterium bovis* Strain”, *Veterinary Microbiology*, v. 70, n. 3-4, pp. 269–276, 1999.