



EDNA LOPES

**SISTEMA NEURAL PARA PREDIÇÃO DE
BRUCELOSE EM REBANHOS BOVINOS**

LAVRAS-MG

2016

EDNA LOPES

**SISTEMA NEURAL PARA PREDIÇÃO DE BRUCELOSE EM
REBANHOS BOVINOS**

Tese apresentada à Universidade Federal de Lavras, como parte das exigências do Programa de Pós-Graduação em Ciências Veterinárias, área de concentração em Ciências Veterinárias, para a obtenção do título de Doutor.

Profa. Dra. Christiane Maria Barcellos Magalhães da Rocha

Orientadora

Prof. Dr. Danton Diego Ferreira

Prof. Dr. Geraldo Márcio da Costa

Coorientadores

LAVRAS - MG

2016

**Ficha catalográfica elaborada pelo Sistema de Geração de Ficha Catalográfica da Biblioteca
Universitária da UFLA, com dados informados pelo(a) próprio(a) autor(a).**

Lopes, Edna.

Sistema neural para predição de brucelose em rebanhos bovinos /
Edna Lopes. – Lavras : UFLA, 2016.
85 p. : il.

Tese(doutorado)–Universidade Federal de Lavras, 2016.

Orientadora: Christiane M. B. Magalhães da Rocha.

Bibliografia.

1. Defesa sanitária animal. 2. Redes Neurais Artificiais. 3.
Diagnóstico. I. Universidade Federal de Lavras. II. Título.

EDNA LOPES

**SISTEMA NEURAL PARA PREDIÇÃO DE BRUCELOSE EM
REBANHOS BOVINOS**

**NEURAL SYSTEM FOR THE PREDICTION OF BRUCELLOSIS IN
CATTLE HERDS**

Tese apresentada à Universidade Federal de Lavras, como parte das exigências do Programa de Pós-Graduação em Ciências Veterinárias, área de concentração em Ciências Veterinárias, para a obtenção do título de Doutor.

APROVADA em 16 de setembro de 2016.

Prof. Dr. Danton Diego Ferreira	UFLA
Prof. Dr. Geraldo Márcio da Costa	UFLA
Prof. Dr. Andrey Pereira Lage	UFMG
Profa. Dra. Elaine Sales Dornelas	UFLA

Profa. Dra. Christiane M. B. Magalhães da Rocha
Orientadora

LAVRAS-MG

2016

Aos meus pais, Edson e Carmen.
Aos meus irmãos, Emerson e Juninho.

DEDICO

AGRADECIMENTOS

A Deus, pela oportunidade de realizar um sonho e por me conceder força em todos os momentos;

À Universidade Federal de Lavras, em especial ao Departamento de Medicina Veterinária, por permitir a realização deste trabalho;

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), pela concessão da bolsa de estudos;

Ao IMA/UFMG/USP, por disponibilizar o banco de dados para o trabalho. Especialmente a Luciana e aos Prof. Andrey e Fernando pelo apoio e parceria.

À professora e eterna orientadora Christiane (Kity), pela confiança, pelas oportunidades e pela amizade. Agradeço também, por todos esses dez anos de contribuição para meu amadurecimento profissional e, não menos importante, meu crescimento pessoal. Muito obrigada, pela parceria pelos conselhos e risadas.

Ao coorientador Geraldo e a profa. Stela, pelo apoio importantíssimo, pelos ensinamentos e pelas oportunidades que me permitiram crescer, não apenas no doutorado, como também durante toda vida acadêmica.

Em especial, agradeço ao querido coorientador e parceiro Danton, que teve um papel fundamental neste trabalho, pelos infinitos ensinamentos compartilhados e pelo apoio incondicional.

Aos velhos amigos do grupo de sanidade animal do LEPI-Laboratório de Epidemiologia: Débora, Fábio, Jonata, Juliana, Denis e Marina, pela parceria nos trabalhos e nas infinitas risadas;

Aos parceiros do Laboratório de Microbiologia que me acompanharam na primeira etapa dessa jornada: Gleí, Juliana Rosa, Marco Túlio e Dircéia, pela dedicação e pelo apoio.

À ADAPAR, que me proporcionou uma experiência profissional ímpar, complementando minha formação na área de defesa animal. Aos queridos colegas de trabalho: Cristina, Beth, Tiago, Deres, Laura, Paulo, Jhonisson, Priscila, Roberto, Elaine, Manoel, Antônio e João Paulo. Em especial, agradeço a Cristina, minha mentora e amiga pelos ensinamentos compartilhados, pela paciência... Ao grande amigo Tiago, pela parceria sem precedentes... Aos amigos Laura e Paulo, por fazerem de minha temporada no Paraná inesquecível e adorável. Muito obrigada, pelos laços de amizade criados e eternizados.

Aos meus queridos pais, pelo amor incondicional, compreensão e paciência em todos os momentos; e aos meus amados irmãos Emerson e Juninho, pela amizade e pelos laços eternos;

Agradeço também a toda a minha família: tios e primos, a minha avó Luiza; e aos meus amigos de Barbacena, em especial, Samira, Bia e Vô Raimundo, muito obrigada por serem meu alicerce e por estarem sempre por perto, mesmo estando longe;

Às amigas de república: Dri, Bruna e Flávia, pelos anos de convivência que deixaram muita saudade e pela companhia tanto nos momentos difíceis quanto naqueles mais alegres;

E a todos aqueles que contribuíram de alguma forma, direta ou indiretamente, para a conclusão deste trabalho. Muito obrigada!

RESUMO

A brucelose é uma das mais importantes doenças na produção de bovinos no Brasil e os programas de defesa animal vigentes ainda encontram dificuldades, principalmente, operacionais, em seu combate. A utilização de ferramentas de inteligência computacional, conhecida como Redes Neurais Artificiais (Redes NA), pode ter grande utilidade nos serviços de vigilância sanitária e epidemiológica para auxílio aos diagnósticos. Esse trabalho teve como objetivo desenvolver um modelo matemático utilizando Redes NA para a classificação de rebanhos positivos e negativos para a Brucelose Bovina, que possa ser utilizado para essa predição. Além disso, foram comparadas formas de tratamento prévio dos dados para melhoria da eficiência do modelo de Rede NA. O desenvolvimento do Sistema Neural compreendeu as etapas de: 1.obtenção dos dados do problema (entrevista com produtores de bovinos); 2.pré-processamento dos dados (seleção de variáveis e análise de qualidade); 3.treinamento das Redes NA (definição da arquitetura da Rede NA); 4.apresentação dos resultados (discriminar o rebanho como positivo ou negativo - modelo de predição). O trabalho foi realizado com dados do inquérito do serviço oficial de defesa sanitária animal do Estado de Minas Gerais (MAPA/ IMA), de setembro de 2010 a dezembro de 2012. A Rede NA foi treinada com diferentes modelos para comparação: 1. com todas as variáveis do banco de dados, sem pré-seleção (49 variáveis) e, 2. com pré-seleções das variáveis de maior relevância. Os métodos de pré-seleção comparados foram: 1. Teste de associação pelo qui-quadrado considerando níveis de significância estatística de 5% ($p < 0,05$) e 10% ($p < 0,10$): dois modelos; 2. Discriminante Linear de Fisher com as 10, 20 ou 30 variáveis mais relevantes: três modelos. O Teste Tukey, com 95% de probabilidade, foi utilizado para a comparação entre os valores de sensibilidade, especificidade e eficiência dos modelos resultantes desses processamentos. Todos os modelos demonstraram os melhores resultados para sensibilidade de 83,33%. Os melhores resultados para especificidade variaram entre 53,89 e 66,60% e para eficiência de 68,61 a 74,96%. Não houve diferença significativa ($p > 0,05$) na eficiência e sensibilidade entre modelos. Quanto à especificidade, o modelo de pré-seleção pelo qui-quadrado (Modelo 2 - $p < 0,10$) atingiu o maior valor, porém a média foi significativamente semelhante a outros ($p > 0,05$). Os resultados demonstram que a aplicação das Redes NA para predição de rebanhos positivos e negativos em veterinária é promissora e a pré-seleção de variáveis para buscar modelos mais simples deve ser utilizada, pois a eficiência se mantém. Essa ferramenta deve ser utilizada na defesa animal para diminuição de custos operacionais e direcionamento de políticas de controle de doenças.

Palavras-chave: Defesa sanitária animal. Redes Neurais Artificiais. Diagnóstico. Minas Gerais. Brasil.

ABSTRACT

Brucellosis is one of the most important diseases in Brazilian cattle rearing, and animal defense programs have difficulty, especially operational, in its control. The use of computer intelligence tools, known as Artificial Neural Networks (ANN) can be of great use in sanitary vigilance and epidemiologic services to aid in the diagnosis. This work had the objective of developing a mathematical model using NAA to classify herds positive and negative for Bovine Brucellosis. In addition, forms of previous data treatments were compared in order to improve the efficiency of the ANN model. The development of the neural system was comprised of the following stages: 1) obtaining the data (interview with cattle farmers); 2) pre-processing of the data (variable selection and quality analysis); 3) ANN training (definition of the ANN architecture); 4) result presentation (discriminate the herd as positive or negative – prediction model). The work was conducted with data from the investigation performed by the official animal sanitary defense service of the state of Minas Gerais, Brazil (MAPA/IMA), from September of 2010 to December of 2012. The ANN was trained with different models for comparison: 1) with all database variables, with no pre-selection (49 variables), and 2) with pre-selection of the most relevant variables. The pre-selection methods compared were: 1) association test by chi-square method, considering the statistical significance levels of 5% ($p < 0.05$) and 10% ($p < 0.10$): row models; 2) Fisher Linear Discriminant with 10, 20 or 30 most relevant variables: three models. The Tukey test with 95% of probability was used for comparison between the sensitivity, specificity and efficiency values of the models resultant of this processing. All models showed the best results for 83.33% of sensitivity. The best results for specificity ranged from 53.89 to 66.60%, and for efficiency, from 68.61 to 74.96%. There was no significant difference ($p > 0.05$) for efficiency and sensitivity between the models. Regarding the specificity, the pre-selection model by chi-square (model 2 – $p < 0.10$) reached the highest value, however, the mean was significantly similar to the other models ($p > 0.05$). The results showed that the application of the ANN for the prediction of positive and negative herds is promising, and the pre-selection of variables to obtain simpler models must be used, given that the efficiency is maintained. This tool must be used for animal defense for reducing operational costs and guidance of disease control policies.

Keywords: Animal sanitary defense. Artificial Neural Networks. Diagnosis. Minas Gerais. Brazil.

LISTA DE FIGURAS

Figura 1 - Estrutura de uma Rede Neural Artificial Perceptron de múltiplas camadas.	30
Figura 2 - Organograma do projeto do Sistema Neural.	38
Figura 3 - Método de seleção das propriedades para compor o projeto do Sistema Neural.	46
Figura 4 - Exemplo do critério de parada do treinamento adotado (<i>early stopping</i>).	47
Figura 5 - Arquitetura das Redes Neurais Artificiais desenvolvidas nesse estudo.	48
Figura 6 - Gráfico sobre a relevância das variáveis para o modelo de Redes Neurais pelo Discriminante Linear de Fisher.	57
Figura 7 - Gráficos de comparação dos valores médios e Intervalos de Confiança (IC 95%) entre os diferentes modelos de pré-processamento utilizados para o desenvolvimento das Redes Neurais Artificiais, Teste de Tukey.	67
Figura 8 - Implementação da Rede Neural Artificial por meio do software Excel.	72

LISTA DE TABELAS

Tabela 1 - Resultados para o desempenho da Rede Neural Artificial (49x10x1).....	51
Tabela 2 - Fatores associados a positividade do rebanho para brucelose bovina, MG, 2010-2012, considerando valor $p < 0,10$	52
Tabela 3 - Fatores associados a positividade do rebanho para brucelose bovina, MG, 2010-2012, considerando valor $p < 0,05$	53
Tabela 4 - Fatores que apresentaram diferença de médias entre os rebanhos positivos e negativos para brucelose bovina, MG, 2010-2012.....	55
Tabela 5 - Resultados para o desempenho da Rede Neural Artificial composta pelas variáveis selecionadas pelo Teste de Qui-quadrado de Pearson.....	56
Tabela 6 - Variáveis mais relevantes de acordo com Discriminante Linear de Fisher, ordenados pelos valores de J, do maior para o menor.....	58
Tabela 7 - Resultados para o desempenho da Rede Neural Artificial composta pelas variáveis selecionadas pelo Discriminante de Fisher.....	59
Tabela 8 - Comparação entre as principais variáveis qualitativas selecionadas pelo Teste de Qui-quadrado de Pearson e Discriminante de Fisher (Continua).....	63
Tabela 9 - Comparação entre a sensibilidade e especificidade das Redes Neurais Artificiais desenvolvidas por meio dos diferentes métodos de pré-seleção.	66

SUMÁRIO

1	INTRODUÇÃO	13
2	OBJETIVOS	17
2.1	Geral	17
2.2	Específicos	17
3	REVISÃO DE LITERATURA	19
3.1	Etiologia da brucelose	19
3.2	Patogenia, Sinais Clínicos e Transmissão da brucelose bovina	20
3.3	Importância econômica da brucelose bovina	22
3.4	Histórico sobre a prevalência da brucelose bovina no Brasil	23
3.5	Diagnóstico e Controle da Brucelose Bovina –PNCEBT	25
3.6	Estudos epidemiológicos observacionais em defesa sanitária animal	27
3.7	Redes Neurais Artificiais (Redes NA)	28
3.7.1	Arquitetura das Redes NA	29
3.7.2	Uso das Redes NA na medicina humana	30
3.7.3	Uso das Redes NA na medicina veterinária	33
4	MATERIAL E MÉTODOS	37
4.1	Dados utilizados: Inquérito de MG, 2010-2012, IMA/MAPA	38
4.2	Análise dos Dados	40
4.2.1	Pré-processamento dos dados	41
4.2.2	Codificação e Normalização	41
4.2.3	Pré-seleções de variáveis para alimentar o modelo de Redes Neurais	42
4.2.4	Algoritmo K-means	43
4.2.5	Projeto da Rede Neural	45
4.2.6	Comparação entre os modelos de Redes Neurais	49
5	RESULTADOS E DISCUSSÃO	51
5.1	Sem pré-processamento – totalidade das variáveis levantadas	51
5.2	Pré-processamento com seleção de variáveis por Análise Univariada	52
5.3	Pré-processamento com seleção de variáveis pelo Discriminante Linear de Fisher	56
5.4	Comparação entre os métodos de pré-processamento utilizados	59
5.5	Redes Neurais Artificiais: arquitetura e desempenho	64
6	CONSIDERAÇÕES FINAIS	73
7	CONCLUSÃO	75
	REFERÊNCIAS	77

1 INTRODUÇÃO

A bovinocultura é uma das mais importantes atividades do agronegócio brasileiro no cenário mundial. O Brasil possui o segundo maior rebanho efetivo do mundo, com cerca de 200 milhões de cabeças, sendo que Minas Gerais, conta com um efetivo de mais de 23 milhões de bovinos (INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA, 2015). Segundo o censo agropecuário de 2006, o valor total da produção de bovinos foi maior que 31 bilhões de reais, em 1.639.765 estabelecimentos existentes da espécie no Brasil (IBGE, 2012). A cadeia produtiva brasileira se destaca nos dois segmentos, carne e leite. No 2º trimestre de 2016, foram abatidas 7,63 milhões de cabeças de bovinos e adquiridos 5,17 bilhões de litros de leite cru por estabelecimento sob algum tipo de inspeção. E Minas Gerais é o primeiro colocado no *raking* dos estados para a produção leiteira e o sexto em abate de bovinos, em 2016 (IBGE, 2016).

A brucelose é uma das mais importantes doenças na produção de bovinos no Brasil. Além de sua importância como zoonose, possui alta prevalência, provoca problemas reprodutivos nos animais, traz grandes prejuízos relacionados às perdas na produção de carne e leite e, principalmente, àqueles relacionados a embargos econômicos (POESTER et al., 2009; SANTOS et al., 2013). Por esses motivos é uma das doenças que está incluída na agenda sanitária brasileira, por meio do Programa Nacional de Controle e Erradicação da Brucelose e da Tuberculose Animal (PNCEBT), implementado pelo Ministério da Agricultura Pecuária e Abastecimento (MAPA), juntamente com as agências estaduais de defesa animal (BRASIL, 2001).

Todavia, os programas de defesa animal, vigentes no Brasil, ainda enfrentam dificuldades no estabelecimento de suas ações. Alguns entraves são observados, principalmente, em decorrência da escassez de recursos humanos e financeiros, somada às características do próprio país, como o grande efetivo de rebanhos, a extensa área territorial e suas diversidades topográficas, climáticas,

culturais e sociais. Dessa forma, métodos que auxiliem na triagem de propriedades com riscos diferenciados para uma doença e, proporcionem melhor direcionamento em seu combate, podem contribuir para a racionalização e eficiência dos programas de sanidade animal.

A utilização de uma ferramenta de inteligência computacional, conhecida como Redes Neurais Artificiais (Redes NA) (HAYKIN, 2008), pode ter grande utilidade nos serviços de vigilância sanitária e epidemiológica para auxílio aos diagnósticos, principalmente por sua boa capacidade de discriminação entre classes, a partir de informações (variáveis) adquiridas do meio (STÄRK; PFEIFFER, 1999).

Segundo Santos et al. (2003), as Redes NA podem ser definidas como modelos não-lineares, podendo ser aplicadas em problemas de regressão, classificação e redução de dados. Além disso, são aplicadas frequentemente em situações onde existem interações não-lineares entre as variáveis dependentes e as independentes.

Tal ferramenta já está implantada em rotinas de diferentes áreas do conhecimento, em razão de sua grande versatilidade na resolução de problemas. Exemplos de uso são: sistemas de reconhecimento de padrões e compressão de voz e imagens, pilotos automáticos de aviões, desenvolvimento de produtos, controle de processos, análises de eletroencefalograma e eletrocardiograma, entre outros. Além disso, pode ser bastante útil na análise de dados epidemiológicos, pelo fato de não haver necessidade de independência e normalidade das variáveis em estudo, e de sua grande capacidade de aprendizado a partir do ambiente.

A discriminação de rebanhos potencialmente positivos para brucelose permite o direcionamento das ações de diagnóstico e controle, minimizando a propagação da doença e exposição dos seres humanos. Dessa forma, o uso de um sistema automático para segregar rebanhos pode dar suporte aos órgãos de

Defesa Sanitária, facilitando a triagem de propriedades. Isso proporciona a otimização dos recursos humanos e financeiros dos órgãos regulamentadores e fiscalizadores, contribuindo com o programa vigente no país (PNCEBT), no controle e erradicação da brucelose animal.

A contribuição, no presente estudo, está no desenvolvimento de uma ferramenta baseada em Redes NA para a classificação e segregação dos rebanhos bovinos, quanto à soroprevalência para brucelose. Além disso, comparar algumas formas de tratamento prévio dos dados, que possam melhorar a eficiência dos modelos de Rede NA em brucelose. Essa ferramenta pode ser apropriada para facilitar o planejamento estratégico das ações do serviço de defesa sanitária animal/ PNCEBT, sendo utilizada como triagem. Para tanto, foram utilizados os dados do inquérito realizado pelo Instituto Mineiro de Agropecuária - IMA/ MAPA em Minas Gerais (MG) entre 2010 e 2012.

2 OBJETIVOS

2.1 Geral

Desenvolver um modelo matemático, utilizando Redes Neurais Artificiais para predição de brucelose em rebanhos bovinos.

2.2 Específicos

- a) Determinar as variáveis (características) mais discriminativas para a predição de brucelose em rebanhos bovinos.
- b) Comparar formas de tratamento prévio dos dados para uso dos modelos de Rede NA em brucelose.

3 REVISÃO DE LITERATURA

Esta revisão aborda alguns aspectos acerca da brucelose bovina, sua importância e suas implicações no contexto brasileiro, principalmente em relação ao programa sanitário vigente, o PNCEBT. Além disso, discorreu sobre os métodos epidemiológicos utilizados nos estudos das ciências veterinárias, focando no modelo de redes neurais.

3.1 Etiologia da brucelose

As espécies do gênero *Brucella* são cocobacilos gram-negativos que medem entre 0,5 até 1,5µm, pleomórficos, aeróbios, não formadores de esporos, não toxigênicos e não fermentadores. Apesar de serem imóveis, apresentam todo o aparato genético para a formação de flagelo, não apresentando, porém, o sistema quimiotático necessário para tal (FRETIN et al., 2005). Esses agentes oxidam nitrito em nitrato e são urease e oxidase positivas (THE WORLD ORGANISATION FOR ANIMAL HEALTH, 2012).

Até 1994, as espécies conhecidas eram *B. abortus*, *B. suis*, *B. melitensis*, *B. canis*, *B. ovis* e *B. neotomae*, com os respectivos hospedeiros preferenciais: bovinos, suínos, ovinos/ caprinos, caninos, ovinos e ratos. Em 1994, foi relatado o isolamento de *B. pinnipedialis* e *B. ceti*, em mamíferos aquáticos e, em 2007, *B. microti* foi descrita em roedores (SCHOLZ et al., 2008). São conhecidos sete biovars de *B. abortus*, três de *B. melitensis* e cinco de *B. suis* (VERGER et al., 1987). Em 2009, uma nova espécie de *Brucella*, a *B. inopinata* foi descoberta apenas em humanos (SCHOLZ et al., 2010). Ainda em 2010, uma cepa diferente de todas as que se conhecia até então, foi nomeada de B02, por sua semelhança genética com *B. inopinata* (TILLER et al., 2010). Filogeneticamente *B. inopinata* está mais próxima da *B. ovis*, no entanto, seu reservatório natural e ecológico permanecem desconhecidos (PAPPAS, 2010). Finalmente, outra linhagem com características típicas de *Brucella*, mas distinta das espécies

atualmente descritas foi recentemente isolada a partir de dois babuínos (SCHLABRITZ-LOUTSEVITCH et al., 2009).

Apesar das espécies de *Brucella* possuírem hospedeiros preferencias, várias delas podem infectar outras espécies animais, que não aquelas de preferência (ACHA; SZYFRES, 2001). Essa diversidade de espécies potencialmente infectantes e de reservatórios naturais para as mesmas tornam a epidemiologia da brucelose humana e animal extremamente complexa.

3.2 Patogenia, Sinais Clínicos e Transmissão da brucelose bovina

A brucelose bovina é uma doença infectocontagiosa de caráter geralmente crônico que tem como agente etiológico principal *Brucella abortus*. A cronicidade da infecção reflete o fato de *Brucella* serem parasitas intracelulares facultativos capazes de sobreviver dentro de células fagocíticas, adaptando-se às condições intracelulares e superando as defesas imunes normais do hospedeiro (BALDWIN; GOENKA, 2006; KÖHLER et al., 2002, 2003). O organismo pode invadir e replicar em uma variedade de células, incluindo células epiteliais, trofoblasto placentário, células dendríticas e macrófagos (GORVEL, 2008). A principal estratégia para a sobrevivência em macrófagos parece ser a inibição da fusão fagossoma-lisossoma (ARENAS et al., 2000; BALDWIN; WINTER, 1994; NAROENI et al., 2001).

Como são gram-negativas, sua superfície lipopolissacarídica induz resposta imunológica intensa. As brucelas têm propensão para a invasão do sistema reticuloendotelial, principalmente no interior dos macrófagos e fagócitos ‘não profissionais’. Ali residem em compartimentos especializados, em pH ácido, e multiplicam-se, sem entretanto, interromper o ciclo celular (GORVEL; MORENO, 2002). *Brucella* inibe a apoptose da célula hospedeira, e assim cria um ambiente para sua sobrevivência e replicação. A resposta imune celular predomina, embora iniba a produção do fator de necrose tumoral alfa (TNF- α).

A expressão de anticorpos serve mais como uma ferramenta de diagnóstico da infecção, tendo pouca efetividade na resposta protetora (DORNELES et al., 2015).

Brucella abortus causa manifestações predominantemente reprodutivas nos animais acometidos, podendo ainda infectar o homem e outras espécies domésticas e silvestres. Nas fêmeas, as manifestações clínicas principais incluem os abortamentos, nascimentos de crias fracas e retenção de placenta, enquanto nos machos pode ocasionar quadros de orquites, epididimite e lesões em glândulas anexas. Essas alterações ocasionam a subfertilidade ou infertilidade em ambos os animais acometidos, o que compromete a produtividade dos rebanhos infectados (BRASIL, 2006).

Esse agente sobrevive por várias horas em ambiente úmido como esterco. Em produtos de concepções como placenta e feto, no solo, ar e água esse organismo pode sobreviver por várias semanas (MEMISH; BALKHY, 2004). A concentração dessas bactérias é alta em fezes, urina, placenta, produtos de concepção abortados e leite (D'ANASTASIO et al., 2011).

A principal fonte de infecção para os bovinos são fetos, envoltórios fetais e descargas vaginais, sendo a transmissão por via transplacentária e sêmen de grande importância. A contaminação fecal dos campos por bezerros alimentados com leite contaminado também pode ser uma fonte relevante para a transmissão da doença. A rota mais comum de invasão ocorre pelo trato gastrointestinal, em decorrência da ingestão de pastagens, forragens e água contaminada com *Brucella*. Além disso, as vacas têm o hábito de lamber membranas fetais, fetos e bezerros recém-nascidos, estes podem conter um grande número de *Brucella*, representando uma importante forma de infecção (ACHA; SZYFRES, 2001).

3.3 Importância econômica da brucelose bovina

A brucelose bovina traz prejuízos econômicos ocasionados pela falência reprodutiva e queda na produção de leite e de carne, agravados pela desvalorização comercial dos animais e dos rebanhos acometidos, aliada à perspectiva de imposição de barreiras sanitárias no âmbito do comércio nacional e internacional (BRASIL, 2006; SANTOS et al., 2013).

Segundo Santos et al. (2013), as perdas decorrentes da brucelose bovina no Brasil, para cada fêmea infectada acima de 24 meses de idade, foram estimadas em R\$ 420,12 para rebanhos de leite e R\$ 226,47 para rebanhos de corte. O prejuízo total estimado foi de, aproximadamente, R\$ 892 milhões (equivalentes a \$ 448 milhões de dólares americanos) ao ano. A cada 1% de variação na prevalência, estima-se a variação de 155 milhões de reais no custo da brucelose bovina no Brasil. Contudo, os prejuízos ocasionados pela brucelose não se restringem às alterações reprodutivas, uma vez que os programas de controle e de vigilância também são bastante onerosos e a doença ainda tem importância zoonótica.

Alves et al. (2015) calcularam as perdas, em decorrência da brucelose de uma forma semelhante, mas consideraram os custos da vacinação em diferentes estratégias de cobertura vacinal: 70, 80 e 90%; em dois estados brasileiros com prevalências diferentes para a doença, Mato Grosso (MT) (com alta prevalência, 10,25%) e São Paulo (SP) (com baixa prevalência, 3,81%). O cálculo foi realizado por meio de dois indicadores econômicos: valor presente líquido – VPL e período de retorno. Quando $VPL > 0$, a estratégia foi considerada vantajosa, ao passo que quando $VPL < 0$, a estratégia não era vantajosa, e quando $NPV = 0$ a estratégia era neutra e, o tempo de retorno levou em consideração o retorno sobre o capital investido. Dessa forma, quando as taxas de vacinação foram mais elevadas, os VPLs também foram mais elevados.

Em SP e MT, foram observadas maiores VPLs (R\$ 351 milhões e R\$ 4,7 bilhões, respectivamente) para a taxa de vacinação de 90% (sob os impactos máximos da doença - R\$ 271 milhões e R\$ 4 bilhões, respectivamente). As taxas de vacinação mais elevadas reduziram o tempo e os custos para ambos os estados atingirem 2% de prevalência da brucelose. Em SP, levariam 12 anos para reduzir a prevalência de brucelose de 3,81% a 2% sob uma taxa de vacinação de 90% e 14 anos sob a taxa de vacinação de 70%, sendo o custo da primeira estratégia 90% do custo da última. Em MT, levariam 22 anos para reduzir a prevalência da brucelose de 10,25% a 2% sob uma taxa de vacinação de 90% e 27 anos, sob taxa de vacinação de 70%, sendo o custo da primeira estratégia 85% do custo da última. Também, observou-se que o período de retorno sobre o capital investido foi menor, com taxas de vacinação mais altas. Em SP, o retorno foi de seis anos e no MT dois anos, para a taxa de 90% de vacinação (ALVES et al., 2015).

3.4 Histórico sobre a prevalência da brucelose bovina no Brasil

Em todo o mundo, animais domésticos e silvestres permanecem como reservatório de *Brucella* entre si e para seres humanos, embora a incidência e prevalência da doença variem muito entre países. A maior atenção é dada à brucelose causada por *B. abortus* (CORBEL, 1997) por seus efeitos danosos para o mercado pecuário mundial. Por esse motivo, muitos países têm adotado medidas rígidas de controle ou erradicação da brucelose nos rebanhos bovinos. Em países que investem pouco em programas sanitários nas áreas de produção de leite e carne, a sua incidência é, geralmente, alta (POESTER; GONÇALVES; LAGE, 2002).

O primeiro caso de brucelose a ser reportado no Brasil foi um caso humano por Gonçalves Carneiro, em 1913 (VERONEZI, 1976) No ano seguinte, a brucelose bovina foi relatada primeiramente no Estado do Rio Grande do Sul,

em 1917, no Ceará e em São Paulo, em 1922. No ano de 1928, Mello e Neiva, isolaram *B. abortus* pela primeira vez no país.

Em 1975, um inquérito nacional realizado para a brucelose bovina apontou uma prevalência média de 7,58%, mas com grandes diferenças entre as diferentes regiões: Nordeste: 2,5%; Sul: 4,0%; Norte: 4,1%; Centro-Oeste: 6,8% e Sudeste: 7,5% (BRASIL, 1977).

Inquéritos soroepidemiológicos, realizados entre os anos de 2001 a 2004, em diferentes estados, apontaram que a distribuição da brucelose é bastante heterogênea no Brasil, verificando-se as prevalências mais elevadas nas regiões Centro-Oeste, Nordeste e Norte do país. No Estado de Goiás, verificou-se a prevalência de 17,5% para propriedades e de 3,0% para animais (ROCHA et al., 2009). No Mato Grosso, a prevalência de focos foi de 41,2% e em animais foi 10,2% (NEGREIROS et al., 2009). E, no Mato Grosso do Sul, a prevalência de focos foi de 41,5% e entre animais variou de 4,5 a 12,6%, de acordo com o circuito pecuário (CHATE et al., 2009). As regiões norte e nordeste do país também apresentam maiores índices de prevalência. No estado de Rondônia, em 2004, as prevalências entre rebanhos e animais foram de 35,2% e 6,2%, respectivamente (VILLAR et al., 2009). Já em um estudo realizado entre 2007 e 2009, no Estado do Maranhão, a prevalência de brucelose estimada entre rebanhos foi de 11,4% e entre animais 2,5% (BORBA et al., 2013).

No Sul e Sudeste, as prevalências são em geral mais baixas, segundo dados dos últimos inquéritos, realizados no período de 2001-2004. No Estado do Espírito Santo, os autores encontraram a prevalência de focos de 9,0% e de animais 3,5% (AZEVEDO et al., 2009). Já, em Santa Catarina, foram observados menores níveis, 0,32% de prevalência de focos e 0,06% entre animais (SIKUSAWA et al., 2009). Em Minas Gerais, a prevalência média foi de 6,0% entre rebanhos e de 1,09% entre animais (GONÇALVES et al., 2009).

3.5 Diagnóstico e Controle da Brucelose Bovina –PNCEBT

A brucelose bovina é uma doença erradicável, como foi demonstrado em diversos países, como no norte da América e da Europa e em países do leste asiático e Oceania, como Japão, Austrália e Nova Zelândia (THE WORLD ORGANISATION FOR ANIMAL HEALTH, 2012). Entretanto, o estabelecimento de um programa de erradicação efetivo, embora contenha medidas genéricas comuns, tais como o diagnóstico e o sacrifício de reagentes, não pode ser visto como um procedimento padrão que se adequa a todas as realidades. Dessa forma, variáveis como a prevalência da doença, características socioeconômicas e culturais são de suma importância para o delineamento do programa de controle e erradicação.

Em 2001, foi implantado no Brasil o Programa Nacional de Controle e Erradicação da Brucelose e da Tuberculose Animal (PNCEBT), por meio da Instrução Normativa 02 de 10 de janeiro de 2001 (BRASIL, 2001). Trata-se de um programa que contempla medidas de adesão compulsória e voluntária, cujo objetivo é padronizar o controle da doença dentro das normas sugeridas pelo Código Zoonosológico Internacional. Como medidas voluntárias, propriedades leiteiras podem adquirir a certificação de propriedades livres, enquanto os rebanhos de corte o *status* de rebanhos controlados. A erradicação da brucelose dessas propriedades é realizada por meio do diagnóstico e sacrifício de animais reagentes, até que se obtenham três testes consecutivos negativos do rebanho. Entre as medidas compulsórias destacam-se a vacinação obrigatória de bezerras contra brucelose, com idade entre três e oito meses, o controle de trânsito de animais destinados à reprodução e o sacrifício ou abate sanitário de animais infectados (BRASIL, 2004).

Para a implementação das medidas efetivas de controle para a brucelose bovina, é imprescindível que se disponham de boas ferramentas de diagnóstico da doença. O PNCEBT definiu como provas oficiais para o diagnóstico da

brucelose no Brasil o Teste do Antígeno Acidificado Tamponado (AAT), o teste do 2-Mercaptoetanol (2-ME) realizado em paralelo com a Soroaglutinação Lenta (SAL) e a reação de Fixação de Complemento (RFC) para o diagnóstico individual e o Teste do Anel em Leite (TAL) como prova de triagem de rebanhos infectados (BRASIL, 2004). O TAL é um método prático, preconizado para se detectar anticorpos anti*Brucella* no leite que tem como vantagens: alta sensibilidade, facilidade de execução, rapidez e baixo custo. É utilizado para rastrear rebanhos infectados em áreas de controle e também na vigilância epidemiológica em áreas livres da enfermidade (SILVA JÚNIOR et al., 2007; THE WORLD ORGANISATION FOR ANIMAL HEALTH, 1996).

A SAL+2-ME é geralmente considerada insatisfatória para os fins de comércio internacional. A RFC é uma técnica de diagnóstico mais específica e padronizada do que a SAL, portanto é o teste de referência para o trânsito internacional de animais (THE WORLD ORGANISATION FOR ANIMAL HEALTH, 2012). Além disso, essa técnica permite maior detecção de animais infectados por longos períodos e menor detecção daqueles vacinados após os oito meses. Porém, é um teste trabalhoso e complexo, que exige mão de obra qualificada e laboratório bem equipado (NIELSEN, 2002).

O teste de AAT (Rosa Bengala) e o Teste de aglutinação em Placa Tamponada são considerados testes adequados para triagem em nível nacional (THE WORLD ORGANISATION FOR ANIMAL HEALTH, 2012). O AAT é uma prova qualitativa que tem como vantagens a alta sensibilidade, praticidade e baixo custo, o que no contexto global resulta em menor custo diagnóstico, possibilitando estratégias de controle da brucelose detectar com maior precocidade animais e rebanhos infectados (MEGID et al., 2000).

O 2-ME é uma prova quantitativa seletiva por meio da detecção de IgG no soro, que é a imunoglobulina indicativa de infecção crônica é realizada em paralelo com a SAL, que detecta, principalmente, IgM (NIELSEN, 2002). Os

dois testes se complementam e podem ser utilizados de forma satisfatória para a confirmação de resultados após uma reação positiva no teste de triagem (PAULIN et al., 2002). A vantagem dessa técnica está no seu baixo custo e facilidade de execução, quando comparado a RFC.

Outra forma alternativa de diagnóstico é o teste de polarização fluorescente (TPF), que permite um diagnóstico rápido e preciso e pode ser realizado fora do laboratório de diagnóstico. A base do teste é pela diferença de rotações de moléculas de tamanhos diferentes, usando para isso um marcador. O TPF pode ser executado por meio de sangue, soro sanguíneo ou leite. É um teste robusto, relativamente barato, simples e muito rápido, já validado para um grande número de espécies, incluindo bovinos, suínos, bisonte e um número de cervídeos (NIELSEN; GALL, 2001)

Visto que a quantidade de testes indiretos disponíveis para o diagnóstico de brucelose é extensa, cada país deve escolher aqueles que melhor se adaptem à sua estratégia, de acordo com suas disponibilidades e características. Sendo assim, essa escolha deve levar em consideração fatores como: custo, o tamanho e as características da população sob vigilância, a situação epidemiológica da doença, a sensibilidade e a especificidade dos testes e a utilização de vacinas (BRASIL, 2006).

Segundo Godfroid et al. (2002) é importante uma combinação de testes e inquéritos epidemiológicos para a erradicação da brucelose, e que, além de resultados de testes, esses fatores são importantes para garantir o sucesso de um programa de erradicação.

3.6 Estudos epidemiológicos observacionais em defesa sanitária animal

Os estudos em epidemiologia são, principalmente, observacionais, uma vez que em várias situações, os pesquisadores não possuem condições de controlar as variáveis independentes. Portanto, na análise epidemiológica, as

variáveis independentes são consideradas fatores de risco caso possam ser associadas às doenças de modo sólido, segundo critérios epidemiológicos (PEREIRA, 2005; ROTHMAN; GREENLAND, 1998; ROUQUAYROL; GURGEL, 2013).

A epidemiologia moderna é uma área que cada vez mais avança nas metodologias de análise de dados coletados em estudos observacionais, fazendo interface com a estatística, a matemática e as ciências sociais. Porém, ainda é uma ciência incompreendida no Brasil por grande parte dos profissionais das áreas médicas e veterinárias, sendo o modelo de epidemiologia descritiva-ecológica (Agente etiológico/hospedeiro/meio ambiente) aquele que prevalece no pensamento hegemônico dessas categorias. Na medicina veterinária, tal avanço pode ser observado em estudos de fatores de risco (ASMARE et al., 2013; BORBA et al., 2013; COWIE et al., 2014) e inclusive de redes complexas (ANKINAKATTE et al., 2013; MCEVOY; AMIGO, 2013; STARK; PFEFFER, 1999).

Adicionalmente, o emprego de análises estatísticas de associação univariada ou modelos múltiplos de regressão logística para avaliação de riscos previamente ao desenvolvimento das redes neurais pode acrescentar informações relevantes à variação conjunta de variáveis respostas. A estatística multivariada é definida como um conjunto de métodos estatísticos utilizados em situações nas quais várias variáveis, contínuas ou não, são medidas simultaneamente em cada unidade experimental (MEDRONHO et al., 2009).

3.7 Redes Neurais Artificiais (Redes NA)

Redes Neurais Artificiais são técnicas computacionais que apresentam um modelo matemático inspirado na estrutura neural de organismos inteligentes e que adquirem conhecimento, por meio da experiência. O cérebro de um mamífero pode ter muitos bilhões de neurônios, já uma grande REDE NA não

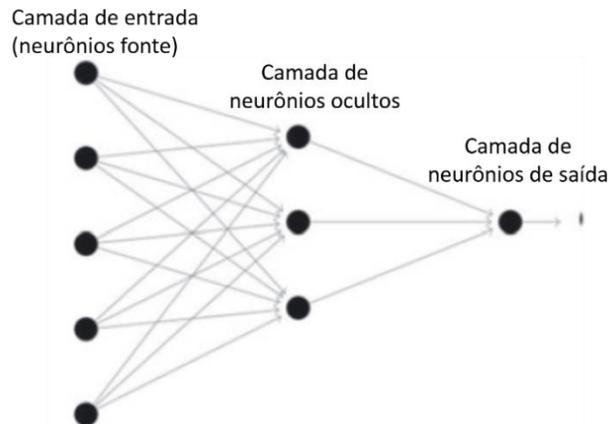
passa de centenas ou milhares de unidades de processamento (BRAGA; CARVALHO; LUDERMIR, 2007).

Três trabalhos foram de extrema relevância para o início dos estudos sobre as Redes NA: McCulloch e Pitts (1943), que introduziram o primeiro modelo de redes neurais simulando “máquinas”; Hebb (1949) descrevendo o modelo básico de rede de auto-organização e Roseblatt (1958), trazendo o modelo tipo Perceptron de aprendizado supervisionado (BRAGA; CARVALHO; LUDERMIR, 2007).

3.7.1 Arquitetura das Redes NA

Das diversas arquiteturas (estruturas) de redes, a mais utilizada em aplicações de reconhecimento de padrões é a rede alimentada diretamente com múltiplas camadas. Conforme definido em Haykin (2008), essa rede consiste de um conjunto de unidades sensoriais que constituem a camada de entrada, uma ou mais camadas ocultas e uma camada de saída. O sinal de entrada se propaga para frente, por meio da rede, camada por camada. Tais redes são, normalmente, chamadas de perceptrons de múltiplas camadas (MLP - *Multilayer Perceptron*). Na Figura 1, ilustra-se um perceptron com três camadas, sendo uma camada oculta.

Figura 1 - Estrutura de uma Rede Neural Artificial Perceptron de múltiplas camadas.



Fonte: Adaptado de Mcevoy e Amigo (2013).

Os perceptrons de múltiplas camadas têm sido aplicados com bastante sucesso para resolver diversos problemas complexos, por meio do seu treinamento de forma supervisionada com um algoritmo muito popular conhecido como algoritmo de retropropagação de erro (*error back-propagation*) (ROSENBLATT, 1962).

Conforme destacado por Santos et al. (2003), a computação envolvida nas etapas de aprendizado da REDE NA é facilitada se o especialista do problema é posto para trabalhar em conjunto com o processamento neural, criando um enfoque de processamento híbrido que ataca o problema.

3.7.2 Uso das Redes NA na medicina humana

Na área médica, as redes neurais artificiais representam um novo paradigma metodológico. Sua aplicação vem se tornando extremamente eficiente e eficaz em inúmeras áreas da medicina, principalmente na área de diagnóstico, prognóstico e terapia. Pelo fato de não haver necessidade de independência e normalidade das variáveis em estudo, bem como a grande

capacidade de aprendizado a partir do ambiente, a aplicação de redes neurais artificiais na análise estatística de dados epidemiológicos tem tido grande aceitação (STARK; PFEFFER, 1999). Além do mais, o processamento neural é capaz de extrair correlações das variáveis de entrada diretamente sobre os espaços de dimensão elevada que tipicamente as caracterizam, tornando tal processamento uma ferramenta valiosa em problemas complexos de reconhecimento de padrões (SANTOS et al., 2005).

As Redes NA vêm sendo utilizadas na medicina humana, principalmente, como ferramenta de auxílio no diagnóstico e prognóstico na clínica médica e análise e predição da ocorrência de doenças (CUNHA et al., 2010; JANE; NEHEMIAH; ARPUTHARA, 2016; SANTOS et al., 2003).

Jane, Nehemiah e Arputhara (2016) utilizaram a Rede NA para auxílio no diagnóstico da doença de Parkinson, e análise da gravidade dos distúrbios em pacientes afetados, utilizando um tipo de algoritmo de treinamento chamado *Q-learning induced backpropagation*. Segundo os autores, os resultados foram satisfatórios em termos de precisão da classificação e demonstrados pela eficiência do algoritmo utilizado.

Cunha et al. (2010), com o intuito de auxiliar os coordenadores de saúde no planejamento e gestão de recursos, estudaram o uso de uma Rede NA para previsão da incidência da malária no município de Cantá, no Estado de Roraima. Os dados utilizados foram coletados no site do Ministério da Saúde, entre 2003 e 2009. A rede foi estruturada com três neurônios na camada de entrada, duas camadas intermediárias e uma camada de saída com um neurônio, e para o treinamento utilizou-se o método *backpropagation*. Segundo os autores, a Rede NA mostrou-se adequada para previsão da doença, determinando com pequenos erros absolutos os valores preditivos, quando comparados ao modelo de regressão logística e aos valores reais.

Também no Brasil, a Rede NA foi estudada para a utilização como ferramenta auxiliar na saúde pública no diagnóstico da Tuberculose Pulmonar (SANTOS et al., 2003). Para tal, 136 pacientes de Unidades de Saúde foram incluídos no estudo, realizado no Rio de Janeiro, entre 2001 e 2002. Os autores, Santos et al. (2003), utilizaram somente sintomas e sinais físicos para construir a rede neural. A REDE NA classificou corretamente 73% dos pacientes da amostra de teste, o que sugere que a REDE NA pode ser uma ferramenta útil para otimizar a utilização de testes mais caros e evitar custos de tratamento desnecessário.

O projeto Neural Tuberculose (TB) é outro exemplo de uso das Redes NA em prol da saúde. Este é uma ação conjunta entre o Laboratório de Processamento de Sinais, do Instituto Alberto Luiz Coimbra de Pós-Graduação e Pesquisa em Engenharia (LPS/COPPE/UFRJ) e a Faculdade de Medicina da Universidade Federal do Rio de Janeiro. O Neural TB é um sistema que oferece uma interface fácil e eficiente com seus usuários. Ao acessar o sistema, que pode ser via Web o usuário preenche um formulário HTML com as informações (sintomas) sobre o paciente a ser analisado (MAIDANTCHIK et al., 2007; SEIXAS et al., 2006). As informações são validadas e, em seguida, alimentam um Sistema Neural, onde as redes *perceptron* multi-camadas (MLP) e ART (*Adaptive Resonance Theory*) (HAYKIN, 2008) são utilizadas. Este, ao ser executado, calcula a probabilidade do paciente ter ou não a tuberculose pulmonar, baseando-se em sintomas e sinais obtidos diretamente do paciente (MAIDANTCHIK et al., 2007; SEIXAS et al., 2006), além de identificar como o paciente se situa no universo de informações disponíveis na base de dados. O sistema serve, portanto, de apoio para tomadas de decisões clínicas mais adequadas ao tratamento de cada doente.

3.7.3 Uso das Redes NA na medicina veterinária

As Redes NA já vêm sendo empregadas ao longo dos anos em estudos veterinários, para diversas espécies, principalmente na área da produção animal. As redes têm sido utilizadas na análise de parâmetros produtivos das espécies e, assim como na medicina humana, como uma ferramenta para auxílio no diagnóstico de enfermidades (AHMADI; GOLIAN, 2011; ANKINAKATTE et al., 2013; MCEVOY; AMIGO, 2013; ROUSH et al., 1996, 1997).

Em 1996, Roush et al. (1996) estudaram a predição de ascite em frangos por meio das Rede NA, e compararam os resultados de diagnósticos de laboratório com incidência predita pela rede neural. Os autores analisaram 15 variáveis fisiológicas das aves e utilizaram uma rede do tipo backpropagation com três camadas (uma camada de entrada, uma camada oculta e uma camada de saída) e, segundo eles, a rede neural identificou corretamente a presença ou ausência de ascite. Em 1997, Roush et al. (1996) utilizaram novamente a Rede NA como uma predição probabilística de ascite em frangos de corte, evitando as intervenções cirúrgicas e outros procedimentos utilizados em 1996. Os resultados encontrados foram promissores para o auxílio na escolha e desenvolvimento de estirpes de corte sem propensão para a ascite, com níveis de sensibilidade de 0,95 (19/20) e especificidade de 0,88 (14/16 aves). Isso que demonstra que a Rede NA é uma ferramenta que pode auxiliar no diagnóstico, reduzindo ou eliminando métodos tradicionais invasivos.

Ainda em aves, Ahmadi e Golian (2011), utilizaram um modelo de Rede NA para avaliar a resposta de pintos de corte e conversão alimentar. Para isso, utilizaram um total de 420 pintos para 28 dietas diferentes. Segundo os autores, a rede apresentou uma precisão elevada para predição da relação entre as concentrações de nutrientes alimentares e desempenho de frangos de corte.

Ankinakatte et al. (2013) e Cavero et al. (2008) utilizaram as Redes NA para detecção precoce de mastite em vacas leiteiras. Ankinakatte et al. (2013)

utilizaram dados de 87 lactações, para compor a Rede NA. O seu desempenho foi avaliado com os dados de validação definidas em termos de sensibilidade e especificidade e encontraram uma sensibilidade de quase 75% para 80% de especificidade fixa. Caverio et al. (2008) utilizaram um conjunto de dados de 403.537 ordenhas envolvendo 478 vacas. Para desenvolver e verificar o modelo, o conjunto de dados foi dividido aleatoriamente em dois subconjuntos, treinamento e teste. Os resultados para os dados, treino e teste, indicaram a generalização do modelo. Porém, os autores encontraram taxas de erro entre 51,3% e 80,5%, que, segundo eles, não proporcionaram um desempenho satisfatório da rede (CAVERO et al., 2008).

Também em vacas, Ghotoorlar et al. (2012) utilizaram a Rede NA com intuito de desenvolver um sistema de pontuação de claudicação automatizado comparável com o método subjetivo convencional. Um grupo de 105 vacas leiteiras e 23 características extraídas a partir de dados da força de reação do solo foram utilizadas. Os autores encontraram sensibilidade e especificidade maiores que 72%, com desvio padrão de 0,31 e o coeficiente médio de variação de 14,55%. Tais resultados, demonstram uma repetibilidade elevada da Rede NA, quando comparada com os métodos de pontuação baseados em visão subjetiva.

Schobesberger e Peham (2002) estudaram também sobre o uso das Redes NA para determinar claudicação, porém em equinos. Para isso, foi usada uma combinação de rastreamento de sinal automático de 175 cavalos em uma escada rolante. A rede foi treinada para distinguir o andar saudável do patológico e quantificar a claudicação. A classificação foi correta em 78,6% dos casos. E, segundo os autores, após treinamento adequado, as redes neurais são potencialmente capazes de fazer um diagnóstico não-humano de claudicação equina.

Não só na área de produção animal, como também em pequenos animais, as Redes NA vêm mostrando-se úteis como uma ferramenta para auxiliar na solução de problemas nos chamados *pets* (GAITA; MILITARU, 2013; MCEVOY; AMIGO, 2013). Gaita e Militaru (2013) utilizaram a Rede NA como uma ferramenta para o auxílio no prognóstico de câncer em cães. Com base em informações clínicas e patológicas e um cenário de terapia, a rede foi treinada. Segundo os autores, a rede demonstrou-se promissora para integrar procedimentos diagnósticos e prognósticos atuais com outros métodos e informações novos para predição do estado clínico de pacientes. Ainda na área de *pets*, Mcevoy e Amigo (2013) estudaram o uso das Redes NA com o objetivo de identificar automaticamente as regiões de imagem que contêm a articulação canina do quadril em radiografias da pelve ventrodorsal. Um conjunto de treinamento de imagens (120 de quadril e 80 de outras regiões) foi usado para treinar a rede para classificar imagens de quadril. O desempenho dos modelos foi avaliado, utilizando-se um conjunto de imagens de teste separado (36 de quadril e 20 de outras áreas). Os valores encontrados foram de 8,9% para erro de classificação, 86% para sensibilidade e 100% para especificidade.

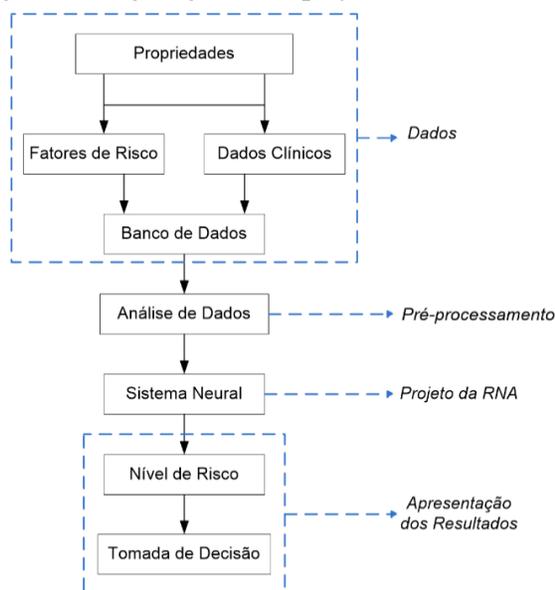
4 MATERIAL E MÉTODOS

Esse estudo aplica modelos de Redes Neurais Artificiais para a classificação de rebanhos como positivos e negativos para brucelose bovina, utilizando os dados do inquérito de MG realizado de 2010 a 2012 pelo IMA/MAPA.

O diagrama em blocos, ilustrado na Figura 2, resume as etapas do desenvolvimento do Sistema Neural, proposto pelo estudo. Os quatro primeiros blocos compreendem a obtenção dos dados do problema (questionários apresentados aos produtores de gado). O próximo bloco é composto por um pré-processamento dos dados (seleção de variáveis relevantes e análise de qualidade). No bloco Sistema Neural, uma rede neural multicamadas é treinada para a detecção da brucelose a partir das informações disponibilizadas. Essa etapa consistiu na definição da arquitetura da Rede NA, ou seja, a definição do número de nodos na camada de entrada, do número de neurônios na camada intermediária, do número de camadas intermediárias, das funções de ativação e do algoritmo de treinamento.

Por fim, os dois últimos blocos compreendem a forma de apresentação dos resultados, de modo a classificar o rebanho do produtor como positivo ou negativo pela brucelose (modelo de predição), de acordo com o enquadramento de suas características.

Figura 2 - Organograma do projeto do Sistema Neural.



4.1 Dados utilizados: Inquérito de MG, 2010-2012, IMA/MAPA

Neste estudo, foram utilizados os dados do inquérito de MG realizado de 2010 a 2012. As informações sobre a coleta dos dados e diagnóstico ocorreram de acordo com o Manual de Procedimentos do estudo epidemiológico sobre Brucelose e Tuberculose em Bovinos e Bubalinos, como descrito por Oliveira et al. (2016). O inquérito foi realizado pela equipe do serviço oficial de defesa sanitária animal do Estado de Minas Gerais (IMA), com o apoio da Coordenação Nacional do PNCEBT e com o suporte do Centro Colaborador do MAPA em saúde animal da FMVZ-USP e FAMV-UnB.

As propriedades sorteadas foram visitadas no período de setembro de 2010 a dezembro de 2012 para preenchimento do formulário (entrevista) e coleta do sangue para o diagnóstico da brucelose, em todo Estado de Minas Gerais

O método de cálculo do número de unidades amostrais baseou-se na estimativa da prevalência de brucelose entre rebanhos. Assim, em cada estrato

regional de amostragem foi realizada uma seleção aleatória de rebanhos. O resultado foi uma prevalência estimada com um intervalo de confiança (IC) de 95%, o qual define a precisão da amostra. Os valores de precisão foram calculados, utilizando-se a fórmula (NOORDHUIZEN et al., 1997): $L = (Z\alpha * SD) / \sqrt{n}$, sendo L = erro absoluto; $Z\alpha$ = valor da distribuição normal para o grau de confiança especificado; SD = desvio padrão = $\sqrt{\text{prevalência} * (1 - \text{prevalência})}$; n = nº de amostras. Assim, no estudo, utilizou-se uma amostra de 300 propriedades/rebanhos (quando possível) ou pelo menos 150 propriedades/rebanho por região do estado.

Dentro das propriedades, selecionaram-se as unidades amostrais secundárias, que foram os animais sob estudo (fêmeas com idade igual ou superior a dois anos). Para isso, foi empregado um método de amostragem aleatória simples ou amostragem aleatória sistemática, dependendo do número de fêmeas com idade igual ou superior a dois anos, existentes na propriedade. Foram selecionadas, de forma aleatória, para exame diagnóstico: 10 fêmeas (≥ 24 meses), quando o rebanho era constituído por até 99 fêmeas (≥ 24 meses), ou todas as fêmeas existentes, nessa faixa etária, quando não totalizavam 10 animais; 15 fêmeas (≥ 24 meses), quando o rebanho era constituído por mais de 99 fêmeas (≥ 24 meses). Foram excluídas das amostras as fêmeas que estavam no período de peri-parto (aproximadamente 15 dias antes do parto e nos 15 dias após o parto).

Amostras sanguíneas foram coletadas por punção da veia jugular, empregando-se uma agulha descartável ou estéril por animal. Foram empregados dois testes, em série, nos soros das fêmeas amostradas. Como teste de triagem, a prova do Antígeno Acidificado Tamponado (AAT). Os soros que reagiram positivamente a este foram submetidos ao segundo teste, confirmatório, o 2-Mercaptoetanol (2-ME). A propriedade foi classificada como positiva, quando houve pelo menos um resultado sorológico positivo de alguma das fêmeas e,

negativa, quando todos os animais apresentavam resultado sorológico negativo para a doença.

4.2 Análise dos Dados

O banco contendo os dados do inquérito de brucelose do Estado de Minas Gerais, realizado conforme seção 4.1 (OLIVEIRA et al., 2016), foi fornecido em forma de planilha do *software* Excel e foi composto pelas variáveis coletadas por meio dos questionários e pelos resultados dos testes de diagnóstico para a doença. O banco de dados foi composto pelos dados de 2.185 rebanhos, dentre eles, 2103 negativos e 82 positivos para brucelose.

Foram consideradas como variáveis dependentes os casos e controles de brucelose (rebanhos positivos e negativos) e como variáveis independentes àquelas coletadas por meio das entrevistas, relacionadas às características e ao manejo geral e sanitário da propriedade. Entre estas, havia variáveis quantitativas, como produção de leite em litros e número de bovinos existentes; e variáveis qualitativas, como tipo de exploração e criação, tipo e número de ordenhas por dia, uso de inseminação artificial, raça bovina predominante, presença de espécies silvestres em vida livre, ocorrência de aborto, manejo de aborto e placenta, realização de testes para brucelose, compra e venda de animais, uso de vacina contra brucelose, abate de bovinos no fim da vida reprodutiva, aluguel de pasto, uso de pastos e/ou aguadas em comum com outras propriedades, acesso dos animais a áreas alagadiças, uso de piquete de parição, local de entrega do leite, resfriamento do leite, assistência veterinária, uso de área para pouso de boiada em trânsito, consumo de leite cru por parte dos proprietários/funcionários da propriedade.

Foram feitas análises descritivas e outras para a validação do banco de dados. Dessa forma, foram descartadas àquelas variáveis que não apresentaram variações ou apresentaram muitos dados perdidos.

4.2.1 Pré-processamento dos dados

Para que o Sistema Neural fosse “alimentado” pelas informações colhidas pelos questionários, realizou-se um pré-processamento dos dados no *Software* SPSS 20.0. Este foi composto por: organização das variáveis e códigos; execução de transformações/ agregações /recodificações das variáveis; verificação de consistência; análise descritiva preliminar; avaliação de *outliers*, seleção das variáveis; análise de associação. Tais técnicas de processamento das informações foram úteis para assegurar a qualidade das mesmas e avaliar a relevância de cada uma delas na detecção da brucelose. Algumas variáveis foram excluídas, em decorrência da inconsistências ou à forte correlação com outras, ou ainda, em razão da baixa relevância como indicativo para a doença, tais como: Vacas (n); Bovinos machos <6 meses; Bovinos machos 6-12 meses (n); Bovinos machos 12-24 meses (n); Bovinos machos >24 meses (n); Bovinos fêmeas 6-12 meses (n); Bovinos fêmeas 12-24 meses (n); Bovinos fêmeas > 24 meses; Produção leite/dia (L); Pousa da boiada no transporte; município. Outras foram recategorizadas para facilitar as análises e o entendimento, como as variáveis: ocorrência de aborto no último ano (sim ou não) e vacinação contra brucelose (sim ou não).

Análises de correlação e associação entre variáveis foram feitas antes da execução da rede. Porém, a opção de manter variáveis nos modelos foi baseada nos resultados de desempenho obtidos no Sistema Neural.

4.2.2 Codificação e Normalização

As variáveis qualitativas foram codificadas atribuindo a elas valores numéricos padronizados de acordo com as respostas dos questionários. As variáveis quantitativas foram normalizadas (padronizadas em uma dada escala numérica) de acordo com a Equação (2):

$$\mathbf{z}_i = \frac{\mathbf{x}_i - \mu_i}{\max(\mathbf{x}_i) - \mu_i}, \quad (2)$$

Em que \mathbf{x}_i representa o vetor com as informações do conjunto de dados referente à variável i e μ_i é a média de \mathbf{x}_i . A normalização dessas variáveis é importante para evitar retardos no aprendizado da rede neural causados pela saturação de neurônios, em decorrência de valores elevados de variáveis.

4.2.3 Pré-seleções de variáveis para alimentar o modelo de Redes Neurais

Foram realizadas pré-seleções das variáveis de maior relevância para a doença, a fim de se obter um modelo mais simplificado, acurado e de fácil aplicação prática. A seguir, serão definidas as pré-seleções utilizadas para comparação posterior.

a. Análise Univariada

Para a pré-seleção, foram realizados estudos de associação das variáveis dependentes (VD) com as variáveis independentes ou explicativas (VI). Os dados qualitativos foram analisados pelo Teste do Qui-quadrado de Pearson, para medida das associações estatísticas. Para as variáveis quantitativas, foram realizadas análises não paramétricas (teste U de Mann-Whitney), pois estas não apresentaram normalidade por meio do teste de Shapiro-Wilk. Os níveis de significância estatística de 5% ($p < 0,05$) e 10% ($p < 0,10$) foram considerados para o desenvolvimento de duas redes neurais diferentes e sua posterior comparação.

b. Fisher

Outro método para redução de variáveis (seleção) foi utilizado a fim de se comparar com o Teste do Qui-quadrado de Pearson, o Discriminante Linear de Fisher (*FDR – Fisher Discriminat Ratio*), bastante usado como ferramenta para pré-seleção de características (DUDA; HART; STORK, 2000). A função do critério FDR é:

$$\mathbf{J}_c = (\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)^2 \odot \frac{1}{\sigma_1^2 + \sigma_2^2}, \quad (3)$$

Onde $\mathbf{J}_c = [J_1 \cdots J_{L_l}]^T$, L_l é o número total de parâmetros, $\boldsymbol{\mu}_1$ e $\boldsymbol{\mu}_2$, e σ_1^2 e σ_2^2 são, respectivamente, os vetores de média e variância dos dados da classe 1 e da classe 2. O símbolo \odot refere-se ao produto de Hadamard $\mathbf{r} \odot \mathbf{s} = [r_0 s_0 \dots r_{L_l-1} s_{L_l-1}]^T$. Pode-se entender que os i -ésimos elementos do vetor de parâmetros, referentes aos maiores valores de \mathbf{J}_c , representam os parâmetros que apresentam melhor separação entre classes. Essa técnica também permite reduzir o número de parâmetros utilizados na classificação e melhorar o desempenho da Rede NA.

4.2.4 Algoritmo K-means

A base de dados foi composta por duas classes, rebanhos positivos para a brucelose e rebanhos negativos. A classe de rebanhos positivos era composta por apenas 82 propriedades, enquanto que a de rebanhos negativos era composta por 2103 propriedades. Observou-se um grande desequilíbrio entre as classes, o que poderia prejudicar o treinamento da Rede NA. Com o objetivo de equilibrar as classes e garantir que o conjunto de treinamento da classe de rebanhos negativos seja composto por propriedades que tragam informações das diversidades da classe, ou seja, que represente bem a classe, as propriedades

negativas foram divididas em 4 agrupamentos. Após a divisão das propriedades negativas em 2, 3, 4, 5 e 6 grupos, escolheu-se a divisão por quatro grupos, em razão da melhor distribuição dos dados nos mesmos, com equilíbrio do número de propriedades em cada grupo. Essa divisão foi alcançada usando o algoritmo k-means. Com o objetivo de formar um banco de dados de treinamento com 130 propriedades negativas, foram selecionadas, aleatoriamente (segundo uma distribuição uniforme), algumas de cada um dos 4 grupos, o que garantiu a estatística equilibrada do grupo de negativos e diminuiu o desvio padrão dos resultados.

K-means é um algoritmo de agrupamento de dados, e foi utilizado para encontrar a melhor divisão de “P” dados em “K” grupos, de maneira que a distância total entre os dados de um grupo e o seu respectivo centro, somada por todos os grupos, seja minimizada.

Após estabelecer o número de clusters previamente definido pelo usuário (4 clusters), o algoritmo calculou os pontos que representam os "centros", chamados centroides. Os centroides iniciais são formados aleatoriamente e posteriormente é calculada a média das distâncias dos vetores de cada grupo aos centroides. Um processo iterativo foi utilizado para encontrar os centroides finais, no qual, em cada passo, os dados são agrupados ao cluster com o centroide mais próximo e posteriormente as médias são recalculadas. O algoritmo converge quando não houver mais alterações nas médias ou quando um número de iterações pré-determinadas for alcançado (THEODORIDIS; KOUTROUMBAS, 2009). O critério de agrupamento do k-means é descrito pela seguinte equação:

$$E = \sum_{k=1}^K \sum_{X_i \in C_k} d(X_i, X_{0k}) \quad (1)$$

Em que K é o número de clusters, X_{ok} é o centroide do cluster C_k e $d(X_i, X_{ok})$ é a distância entre os pontos X_i e X_{ok} . Geralmente, utiliza-se a distância Euclidiana, embora outras distâncias possam ser usadas.

4.2.5 Projeto da Rede Neural

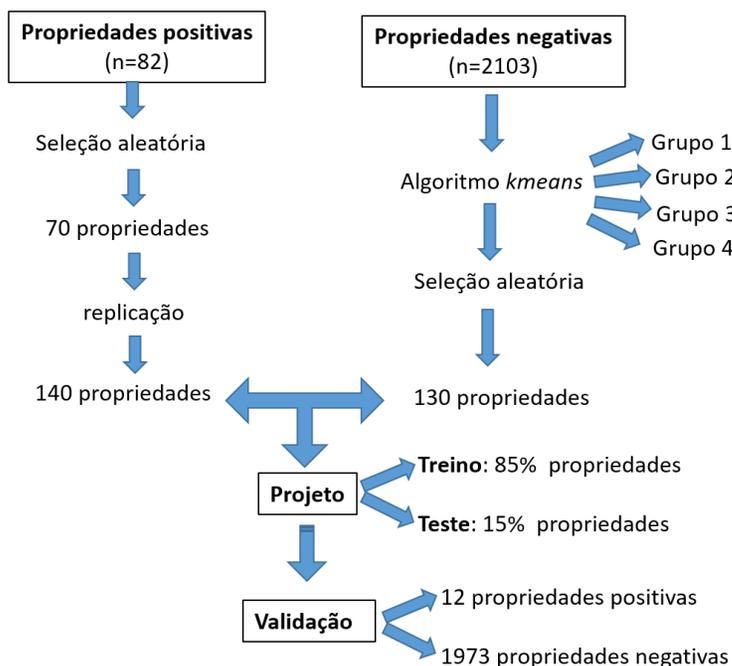
Neste estudo, foram utilizadas redes neurais multicamadas (*Multi Layer Perceptron*) (HAYKIN, 2008), tipicamente composta de camadas alinhadas de neurônios. Nesse tipo de rede, as informações de entrada se propagam adiante, camada por camada, até alcançar a camada de saída.

Vários modelos de redes neurais com uma camada intermediária foram gerados, variando o número de neurônios e adequando o algoritmo de treinamento (*Scaled conjugate gradient backpropagation algorithm*) (MOLLER, 1993). Esse algoritmo foi escolhido por ser mais rápido no treinamento e por apresentar bons resultados para o banco de dados utilizado. O ajuste da rede e o número de neurônios foram selecionados, por meio do cálculo da sensibilidade e especificidade. Assim, optou-se sempre por redes com maior sensibilidade, sem prejuízos importantes em sua especificidade.

A cada execução, os dados foram sorteados e divididos em dois conjuntos, projeto (treino e teste) e validação. Para o projeto, foram escolhidos 70 eventos positivos e estes foram replicados, totalizando 140, na tentativa de contornar ou minimizar os efeitos do mal da dimensionalidade (poucos dados para muitas variáveis) (CASTRO; BRAGA, 2011). Para a classe negativa foram escolhidos aleatoriamente 130 eventos para projeto (de acordo com o algoritmo k-means), a fim de igualar as classes e não favorecer nenhuma delas no processo de treinamento. No projeto da rede neural, 85% das propriedades foram usadas para treino e 15% para teste. O restante dos dados foi usado para validação, ou seja, 12 propriedades para o grupo positivo e 1973 para o grupo negativo. Na figura 3, apresentam-se, esquematicamente, os procedimentos realizados para a

seleção das propriedades negativas e positivas nas fases de projeto e validação da rede neural proposta.

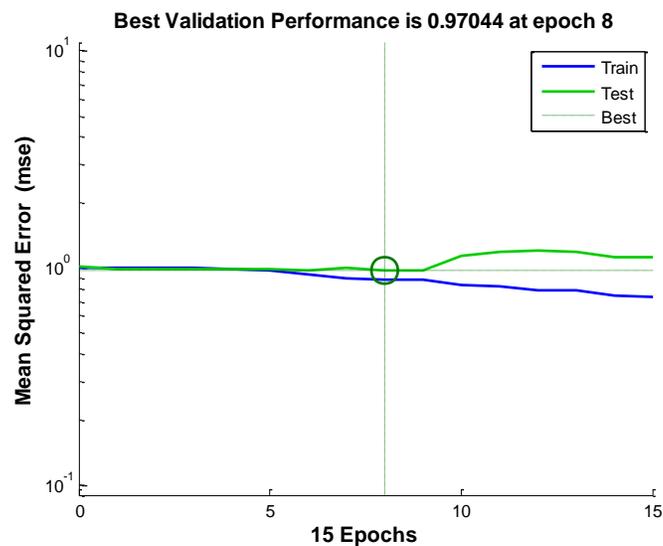
Figura 3 - Método de seleção das propriedades para compor o projeto do Sistema Neural.



O critério de parada utilizado foi o *early stopping*, em que um número máximo de 25 épocas foi utilizado e 15% dos eventos de projeto foram utilizados para o teste (PRECHELT, 1998). Na Figura 4, ilustra-se o *early stopping* em uma das execuções do processo de treinamento da rede neural. Observe que, na oitava época, o erro do conjunto de treino continuou caindo e o do conjunto de teste se manteve e aumentou próximo da décima época. Como o erro do conjunto de teste continuou aumentando e o do conjunto de treino caindo até a décima quinta época, o algoritmo de treinamento parou o treinamento (parada antecipada – *early stopping*) antes de atingir as 25 épocas, e a rede

neural resultante desse treinamento é aquela com os pesos referentes à oitava época, indicada, na figura 4, pela marcação circular, que é a época em que a rede classificou melhor os dados do conjunto de teste. Com isso, o *overfitting* (treinamento em excesso) é evitado e se garante a generalização da rede neural, que é a capacidade de acertar a classificação de novas amostras, não utilizadas no treinamento.

Figura 4 - Exemplo do critério de parada do treinamento adotado (*early stopping*).



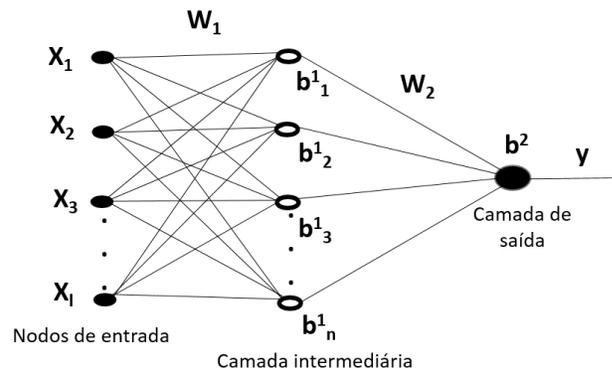
A função de ativação adotada foi a tangente hiperbólica, que implementa a seguinte equação:

$$\tanh(t) = \frac{e^t - e^{-t}}{e^t + e^{-t}} \quad (4)$$

Na figura 5, pode-se observar a estrutura da rede neural desenvolvida nesse trabalho. A metodologia de treinamento e validação da rede neural foi

realizada 100 vezes e, portanto, 100 redes neurais foram obtidas. Os resultados de desempenho no projeto (treino e teste) e na validação foram apresentados na forma de média \pm desvio padrão.

Figura 5 - Arquitetura das Redes Neurais Artificiais desenvolvidas nesse estudo.



Legenda: X_i = variáveis fornecidas às redes (Ex. tamanho do rebanho, número de ordenhas, tipo de exploração, etc); o vetor $\mathbf{X} = [X_1, X_2, \dots, X_i]$ representa uma propriedade;
 \mathbf{W}_1 = matriz de pesos da camada 1 (de entrada);
 \mathbf{W}_2 = matriz de pesos da camada 2;
 b^i_j ; bias, constantes definidas durante o treinamento, assim como \mathbf{W} ;
 $\tanh(\mathbf{X}^T \mathbf{W}_1 + \mathbf{b}^1) =$ saída da camada intermediária (Z);
 $\tanh(\mathbf{Z}^T \mathbf{W}_2 + \mathbf{b}^2) =$ saída da rede (y), se $y > 0 =$ brucelose e se $y < 0 =$ não brucelose.

O número de neurônios na camada intermediária foi definido empiricamente. Adotou-se um único neurônio na camada de saída, tal que valores maiores do que zero indicam propriedades positivas para a Brucelose e valores menores do que zero indicam propriedades negativas.

As Redes NA foram projetadas e testadas usando o *software* MatLab, que apresenta uma biblioteca bem completa para treinamento e testes de redes neurais

4.2.6 Comparação entre os modelos de Redes Neurais

A Rede NA foi primeiramente treinada com todas as variáveis do banco de dados, sem utilização de uma pré-seleção. A camada de entrada da rede foi composta por 61 variáveis. Em seguida, foram utilizados os dados resultantes dos pré-processamentos das análises univariada e do discriminante de Fisher com o objetivo de se comparar essas diferentes abordagens. Assim, o Teste Tukey, com 95% de probabilidade, foi utilizado para a comparação entre os valores de sensibilidade, especificidade e eficiência das Redes NA, resultantes de cada pré-processamento.

Para o cálculo, as seguintes equações foram utilizadas: Sensibilidade = $VP / (VP + FN)$, onde VP = Verdadeiros positivos; FN = Falsos Negativos; Especificidade = $VN / (VN + FP)$, onde VN = Verdadeiros negativos; FP = Falsos Positivos e; Eficiência, calculada pela média dos níveis de sensibilidade e especificidade.

5 RESULTADOS E DISCUSSÃO

Para a obtenção de uma REDE NA mais eficiente foi necessário a seleção aleatória das observações para compor os conjuntos de treino e validação da rede. A ponderação de observações pode ser aplicada levando em consideração a quantidade desequilibrada de observações em diferentes classes (STARK; PFEIFFER, 1999).

Os resultados das Redes NA a seguir foram subdivididos em tópicos de acordo com o tipo de pré-processamento utilizado.

5.1 Sem pré-processamento – totalidade das variáveis levantadas

A Rede NA foi alimentada por todas as 49 variáveis do conjunto de dados e os melhores resultados foram obtidos para 10 neurônios na camada intermediária, o que levou à arquitetura 49x10x1, ou seja, 49 neurônios na camada de entrada (dados inseridos para alimentar a rede), 10 neurônios na camada intermediária (camada oculta) e um neurônio na camada de saída (resultado: rebanho positivo ou negativo). Na Tabela 1, apresenta-se a média de desempenho da rede no projeto (treino e teste) e na validação, bem como o melhor resultado de desempenho encontrado dentre as 100 execuções realizadas.

Tabela 1 - Resultados para o desempenho da Rede Neural Artificial (49x10x1).

Rebanhos	Treino +teste ± DP (%)	Validação ± DP (%)	Melhor resultado Validação (%)
Positivos	81,14 ± 14,97	53,58 ± 17,50	83,33
Negativos	77,93 ± 15,16	57,93 ± 10,15	61,51
Média*	-----	-----	72,42

* Média dos melhores resultados de validação (rebanhos positivos e negativos)

5.2 Pré-processamento com seleção de variáveis por Análise Univariada

As variáveis foram selecionadas para camada de entrada da rede por duas diferentes formas. Para a primeira pré-seleção das variáveis, considerou-se $p < 0,10$ no Teste de Qui-quadrado de Pearson, totalizando 21 variáveis para o desenvolvimento do Sistema Neural, dentre elas, variáveis qualitativas nas tabelas 2 e 3 e variáveis quantitativas na Tabela 4.

Na segunda pré-seleção, as variáveis associadas foram selecionadas pelo mesmo teste, considerando $p < 0,05$, totalizando 13 variáveis para o desenvolvimento do Sistema Neural. Na Tabela 3, foram apresentadas as variáveis qualitativas e na Tabela 4, as variáveis quantitativas.

Tabela 2 - Fatores associados a positividade do rebanho para brucelose bovina, MG, 2010-2012, considerando valor $p < 0,10$

(Continua)

Característica	Categoria	Casos n (%)	Controles n (%)	Valor P ¹
Raças bovinas	Zebu	17 (20,7)	308 (14,7)	0,056
	Europeu de leite	11 (13,4)	161 (7,7)	
	Europeu corte	0 (0,0)	1 (0,0)	
	Mestiço	42 (51,2)	1402 (66,8)	
	Outras raças	12 (14,6)	227 (10,8)	
Destino aborto	Não faz nada	39 (53,4)	1244 (66,1)	0,067
	Alimenta porco/cão	1 (1,4)	33 (1,8)	
CR leilão	Enterra/joga fora/queima	33 (45,2)	606 (32,2)	0,058
	Não	77 (93,9)	2048 (97,4)	
VR fazendas	Sim	5 (6,1)	55 (2,6)	0,09
	Não	64 (78,0)	1786 (84,9)	
Aluga pasto	Sim	18 (22,0)	317 (15,1)	0,095
	Não	56 (68,3)	1605 (76,3)	
	Sim	26 (31,7)	498 (27,3)	

Tabela 2 - Fatores associados a positividade do rebanho para brucelose bovina, MG, 2010-2012, considerando valor $p < 0,10$

(Conclusão)

Característica	Categoria	Casos n (%)	Controles n (%)	Valor P ¹
Destino leite	Não entrega	28 (34,1)	888 (42,2)	0,096
	Cooperativa	19 (23,2)	385 (18,3)	
	Laticínio	35 (42,7)	751 (35,7)	
	Direto ao consumidor	0 (0,0)	78 (3,7)	
Resfria leite	Não	42 (51,2)	1290 (61,4)	0,064
	Sim	40 (48,8)	812 (38,6)	
Fabrica queijo/ manteiga	Não	63 (76,8)	1429 (68,0)	0,092
	Sim	19 (23,2)	2,0)	

¹ Teste de Qui-quadrado de Pearson - análise univariada.

Legenda: CR = compra reprodutor; VR= vende reprodutor.

Tabela 3 - Fatores associados a positividade do rebanho para brucelose bovina, MG, 2010-2012, considerando valor $p < 0,05$

(Continua)

Característica	Categoria	Casos n (%)	Controles n (%)	Valor P ¹
Tipo de ordenha	Não ordenha	18 (22,0)	417 (19,9)	0,017
	Manual	40 (48,8)	1327 (63,2)	
	Mecânica ao pé	15 (18,3)	240 (11,4)	
	Mecânica sala de ordenha	9 (11,0)	115 (5,5)	
Inseminação dicotômico	Não	64 (78,0)	1827 (87,0)	0,019
	Sim	18 (22,0)	272 (13,0)	
Presença de Ovinos/caprinos	Não	72 (87,8)	1980 (94,2)	0,018
	Sim	10 (12,2)	123 (5,8)	

Tabela 3 - Fatores associados a positividade do rebanho para brucelose bovina, MG, 2010-2012, considerando valor $p < 0,05$

(Conclusão)

Característica	Categoria	Casos n (%)	Controles n (%)	Valor P ¹
Teste de Brucelose	Não	60 (73,2)	1837 (87,4)	0,000
	Sim	22 (26,8)	265 (12,6)	
CR	Não	39 (47,6)	1254 (59,7)	0,028
	Sim	43 (52,4)	847 (40,3)	
VR	Não	55 (67,1)	1637 (77,9)	0,022
	Sim	27 (32,9)	465 (22,1)	
VR leilões	Não	76 (92,7)	2061 (98,0)	0,001
	Sim	6 (7,3)	42 (2,0)	
Vacina Brucelose dicotômico	Não	1 (1,2)	153 (7,3)	0,027
	Sim	81 (98,8)	1950 (92,7)	
Local de abate de reprodutores	Não abate	30 (36,6)	1187 (56,5)	0,001
	Estabelecimento com inspeção	31 (37,8)	457 (21,8)	
	Estabelecimento sem inspeção	15 (18,3)	269 (12,8)	
	Própria fazenda	6 (7,3)	187 (8,9)	
Piquete de Parição	Não	36 (43,9)	1355 (64,5)	0,000
	Sim	46 (56,1)	745 (35,5)	
Consome leite cru	Não	45 (55,6)	1389 (66,1)	0,049
	Sim	36 (44,4)	711 (33,9)	
Assistência Veterinária	Não	47 (58,0)	1557 (74,1)	0,001
	Sim	34 (42,0)	545 (25,9)	

¹ Teste de Qui-quadrado de Pearson - análise univariada.

Legenda: CR = compra reprodutor; VR= vende reprodutor.

Tabela 4 - Fatores que apresentaram diferença de médias entre os rebanhos positivos e negativos para brucelose bovina, MG, 2010-2012.

	Casos*			Controles*			Valor p ¹
	Mediana	Quartis (q1-q3)	Mín-Máx	Mediana	Quartis (q1-q3)	Mín-Máx	
Vacas (n) **	12,0	2,75-25,75	0-124	7,0	1,0-19,0	0-260	0,024*
Produção leite/dia (L)**	70,0	16,0-212,5	0-2200	30,0	5,0-115,0	0-3300	0,009*
Bovinos machos <6 meses**	2,0	0,0-10,0	0-40	1,0	0,0-4,0	0-120	0,039*
Bovinos machos 6-12 meses (n) **	3,0	0,0-15,25	0-133	2,0	0,0-5,0	0-310	0,004*
Bovinos machos 12-24 meses (n) **	0,5	0,0-12	0-142	0,0	0,0-4,0	0-3353	0,053
Bovinos machos >24 meses (n) **	1,0	0,0-9,0	0-126	1,0	0,0-2,0	0-1121	0,084
Bovinos fêmeas 6-12 meses (n) **	5,0	1,0-11,25	0-103	3,0	0,0-7,0	0-213	0,002*
Bovinos fêmeas 12-24 meses (n) **	8,5	2,0-24,0	0-196	3,0	0,0-10,0	0-291	0,000*
Bovinos fêmeas >24 meses**	31,5	17,25-89,75	2-531	18,0	7,0-41,0	1-758	0,000*
Tamanho rebanho (n)	74,5	30,75-249,25	4-814	40,0	18,0-90,0	1-4541	0,000*

¹Teste U de Mann-Whitney considerando $p < 0,10$; * $p < 0,05$; ** variáveis excluídas do Sistema Neural por estarem correlacionadas entre si, ou por inconsistência dos dados (produção de leite).

Na Tabela 5, apresenta-se a média de desempenho da rede para os dados de projeto (treino +teste) e de validação, bem como os resultados da rede que obteve o maior desempenho dentre as 100 execuções. Para ambas as redes, com 21 e 13 variáveis, os melhores resultados foram encontrados para 20 e 10 neurônios na camada intermediária e as redes neurais foram projetadas com a arquitetura 21x20x1 e 13x10x1, respectivamente.

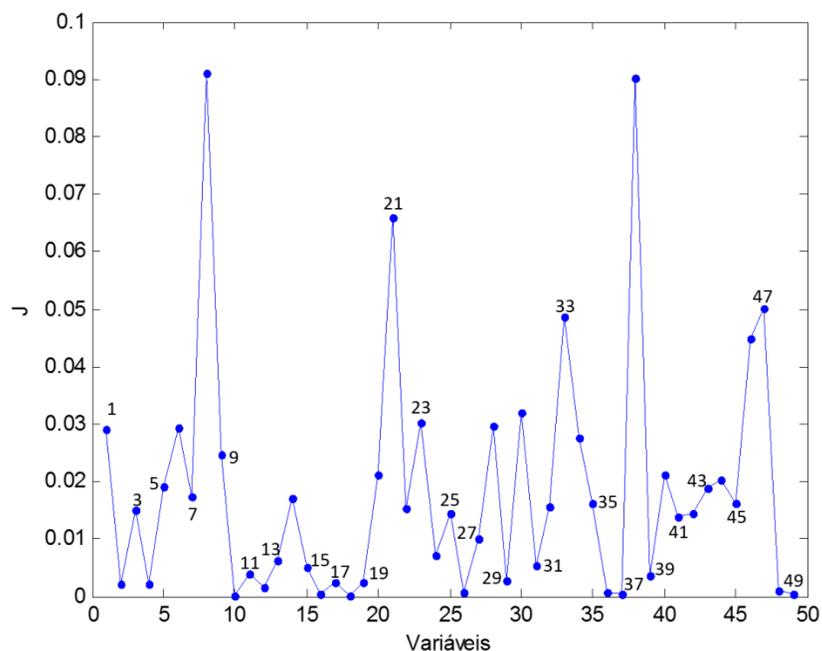
Tabela 5 - Resultados para o desempenho da Rede Neural Artificial composta pelas variáveis selecionadas pelo Teste de Qui-quadrado de Pearson.

Tipo de pré-seleção	Rebanhos	Treino +teste ± DP (%)	Validação ± DP (%)	Melhor resultado Validação (%)
Seleção de 21 Variáveis (21x20x1)	Positivos	77 ± 13,01	52,08 ± 16,21	83,33
	Negativos	75,32 ± 11,18	56,18 ± 8,62	66,60
	Média	-----	-----	74,96
Seleção de 13 Variáveis (13x10x1)	Positivos	71,46 ± 13,85	54,17 ± 18,37	83,33
	Negativos	68,29 ± 9,07	54,42 ± 11,90	59,53
	Média	-----	-----	71,43

5.3 Pré-processamento com seleção de variáveis pelo Discriminante Linear de Fisher

O Discriminante Linear de Fisher indicou as variáveis com maior e menor relevância para o modelo, quanto maior valor de J, maior sua capacidade de separação de classes. Assim, as variáveis mais relevantes são representadas pelos picos mais altos e as menos relevantes pelos pontos mais baixos no gráfico (Figura 6).

Figura 6 - Gráfico sobre a relevância das variáveis para o modelo de Redes Neurais pelo Discriminante Linear de Fisher.



Legenda: J= nível de separação entre classes; 1.região; 2.exploracao; 3.criacao; 4.número ordenha; 5.tipo de ordenha; 6.inseminação dicotômico; 7.raça bovinos; 8.total animais; 9.ovinos e caprinos; 10.equinos; 11.suinos; 12.aves; 13.cão; 14.gato; 15.silvestres; 16.cervideos; 17.capivaras; 18.outro silvestres; 19.aborto dicotômico; 20.destino aborto; 21.teste de brucelose; 22.regularidade teste brucelose; 23.compra reprodutor (CR); 24. CR exposição; 25. CR leilão; 26. CR comerciante; 27. CR fazenda; 28.vende reprodutor (VR); 29. VR exposição; 30. VR leilão; 31. VR comerciante; 32. VR fazenda; 33.vacina brucelose dicotômico; 34.abate reprodutor; 35.aluga pasto; 36.pasto em comum; 37.presença de áreas alagadas; 38.piquete parição; 39.destino leite; 40.resfria leite; 41.como resfria leite; 42.entrega leite a granel; 43.faz queijo manteiga; 44.finalidade queijo manteiga; 45.consome leite cru; 46.assistência veterinária; 47.tipo de assistência; 48.aguadas e bebedouro; 49.classificação da propriedade.

Na Tabela 6, apresentam-se as variáveis mais relevantes pelo Discriminante de Fisher. Assim, três formas de pré-processamento foram utilizadas para compor cada um dos três modelos de redes. Na primeira forma, selecionaram-se as 30 variáveis para camada de entrada da rede (maiores valores

de J). Na segunda, as 20 variáveis mais relevantes foram selecionadas e, por último, as 10 variáveis de maior relevância foram selecionadas pelo Discriminante de Fisher para compor a rede.

Tabela 6 - Variáveis mais relevantes de acordo com Discriminante Linear de Fisher, ordenados pelos valores de J, do maior para o menor.

Variáveis	Identificação no gráfico (fig.4)	Valores de J
Tamanho rebanho (n)	8	0,0909
Piquete parição	38	0,0901
Teste de brucelose	21	0,0658
Tipo de assistência veterinária	47	0,0503
Vacina brucelose dicotômico	33	0,0487
Assistência veterinária	46	0,0448
VR leilão	30	0,0321
CR	23	0,0304
VR	28	0,0298
Inseminação	6	0,0294
Região	1	0,0292
Local de abate de reprodutores	34	0,0276
Ovinos e caprinos	9	0,0246
Resfria leite	40	0,0213
Destino aborto	20	0,0212
Finalidade queijo manteiga	44	0,0202
Tipo de ordenha	5	0,0192
Fabrica queijo/ manteiga	43	0,0189
Raça bovinos	7	0,0175
Gato	14	0,0172
Consome leite cru	45	0,0163
Aluguel de pasto	35	0,0161
VR fazenda	32	0,0157
Regularidade do teste de brucelose	22	0,0153
Tipo de criação	3	0,0150
VR leilão	25	0,0145
Entrega leite a granel	42	0,0144
Como resfria leite	41	0,0140
CR de fazenda	27	0,0100
CR em exposição	24	0,0070

Legenda: CR =Compra reprodutor; VR=Vende reprodutor

Na Tabela 7, apresenta-se a média de desempenho da rede no projeto (treino + teste), validação e, os melhores resultados de desempenho da rede. Para as redes geradas, com 30, 20 e 10 variáveis, os melhores resultados encontrados foram 30, 20 e 10 neurônios na camada intermediária, resultando nas arquiteturas 30x30x1, 20x20x1 e 10x10x1, respectivamente.

Tabela 7 - Resultados para o desempenho da Rede Neural Artificial composta pelas variáveis selecionadas pelo Discriminante de Fisher.

Tipo de pré-seleção	Rebanhos	Treino +teste ± DP (%)	Validação ± DP (%)	Melhor resultado Validação (%)
Seleção de 30 Variáveis (30x30x1)	Positivos	78,40 ± 17,38	51,17 ± 17,37	83,33
	Negativos	77,61 ± 14,91	59,83 ± 10,28	61,51
	Média	-----	-----	72,42
Seleção de 20 Variáveis (20x20x1)	Positivos	74,98 ± 14,45	50,42 ± 15,18	83,33
	Negativos	74,47 ± 12,56	57,88 ± 10,02	53,89
	Média	-----	-----	68,61
Seleção de 10 Variáveis (10x10x1)	Positivos	69,00 ± 12,53	56,75 ± 17,52	83,33
	Negativos	64,82 ± 9,90	51,31 ± 14,74	61,92
	Média	-----	-----	72,62

5.4 Comparação entre os métodos de pré-processamento utilizados

Um dos parâmetros que mais contribuíram para o modelo foi aquele relacionado ao tamanho do rebanho (Gráfico Fisher [GF]= 8, J=0,0909; Análise univariada [AU]: p=0,000 e coeficiente correlação [r]=0,097). Oliveira et al. (2016) utilizaram o mesmo banco de dados do presente trabalho e também encontrou diferença com relação ao número de animais no rebanho, em que rebanhos com mais de 210 animais tinham quase oito vezes mais chance de infecção pela doença. Calistri et al. (2013) e Sanogo et al. (2012) também verificaram que o aumento da infecção foi associado a maiores rebanhos. Os

autores verificaram o aumento do risco para a doença em rebanhos com um maior número de animais, sugerindo que uma maior densidade populacional desses rebanhos aumenta a probabilidade de mais contato entre os animais, além de se aumentar as chances da disseminação do agente no ambiente, em razão do maior número de partições e abortos.

Fatores relacionados ao manejo sanitário do rebanho também apresentaram-se relevantes como a realização de testes de brucelose (GF=21, J=0,0658; AU: p=0,000) e assistência veterinária (GF=46, J=0,0448; p=0,001), vacinação contra a brucelose (GF=33, J=0,0487; AU: p=0,027). Borba et al. (2013), encontraram associação entre a doença na propriedade com a assistência veterinária e a vacinação das bezerras contra brucelose no Estado do Maranhão. Além dessas, outras variáveis relacionadas ao manejo geral da propriedade foram importantes, como a utilização de piquete partição (GF=38, J=0,0901; AU: p=0,000) (Tabela 8). A presença de currais para partição nas propriedades foi também ligeiramente associada ao aumento de rebanhos soropositivos (ASMARE et al., 2013).

Verificou-se que fatores relacionados a introdução de animais foram relevantes para o modelos de Redes NA. A compra e venda de reprodutores apresentaram razoáveis níveis de separação de classes (GF=23 e 28, J≈0,03) e intensidades de associação na análise univariada (p= 0,001; OR=1,632; IC=1,049-2,540 e OR=1,728; IC=1,078-2,771, respectivamente) (Tabela 8). Asmare et al. (2013) verificaram o aumento da soroprevalência nos bovinos, em decorrência da aquisição de animais de outros rebanhos. Cowie et al. (2014) e Gonçalves et al. (2009) identificaram também a compra de animais como sendo um fator de risco para a doença. Segundo Gonçalves et al. (2009), o maior problema não é a aquisição desses animais, mas sim a sua introdução sem a prévia realização de testes ou sem o conhecimento da condição sanitária do rebanho de origem.

A presença de ovinos e caprinos na propriedade apesar de apresentar um menor nível de separação de classes (GF= 20, J= 0,0246), apresentou relevante intensidade de associação (p=0,018 e OR=2,236; IC=1,126-4,440) (Tabela 8). No estudo de Asmare et al. (2013), a presença de pequenos ruminantes influenciou no aumento da prevalência da soropositividade nos rebanhos estudados. Tal fato pode ser decorrente de os animais serem potenciais reservatórios do agente, visto que a infecção de ovinos e caprinos por *B. abortus* já foi relatada na literatura (OCHOLI et al., 2004).

As variáveis que menos contribuíram para o modelo foram observadas no Método de Fisher, pela baixa capacidade de separação de classes (J≈0,00). Tal observação foi ratificada na análise univariada, na qual essas variáveis apresentaram um valor de $p > 0,10$ (dados não apresentados). As variáveis de menor relevância refletem principalmente sobre as características da propriedade e manejo geral dos animais (rotina). Características tais como, bovinos machos castrados (GF=10), presença de equinos (GF=21), aves (GF=23), cervídeos (GF=27) e outros animais silvestres (GF=29), presença de áreas alagadas (GF=48), aguadas e bebedouro (GF=59) e classificação da propriedade (área rural, peri-urbana ou assentamento) (GF=61).

Apesar das variáveis em questão não contribuírem para o modelo de redes e não apresentarem associação com a brucelose, verificou-se, na literatura, a importância de algumas delas para a doença, por isso foram levantadas no inquérito. Borba et al. (2013), destacam a presença de zonas úmidas (pântanos, riachos e rios) como fator de risco por aumentar a concentração dos animais em torno dessas fontes de água, proporcionando a oportunidade de contato próximo e assim a propagação da doença, e em razão também da probabilidade de aumentar a sobrevivência de *Brucella* no ambiente.

Cervídeos e outros ungulados silvestres já foram descritos como reservatórios de *B. abortus* para bovinos (PAULIN; FERREIRA NETO, 2003).

No estudo de Gonçalves et al. (2009), em Minas Gerais, verificou-se uma associação de focos da doença com a presença de cervídeos. Mas, segundo os autores, ainda é necessário caracterizar melhor essa relação, a forma de contato entre as espécies e a importância dos cervídeos como reservatórios da brucelose bovina.

Sobre manejo de rotina, outras variáveis também foram excluídas em ambos os métodos de pré-seleção, Análise Univariada e Discriminante de Fisher, em razão de sua baixa relevância para o desenvolvimento do Sistema Neural, tais como: tipo de exploração (corte, leite ou misto) (GF=2), número de ordenhas realizadas por dia (GF=4), se compra gado de comerciante (GF=37), vende animais para exposição (GF=40), utiliza pasto em comum (Figura 6). Apesar das variáveis relacionadas a comercialização de animais serem mencionadas na literatura como potenciais fatores de risco (ASMARE et al., 2013; COWIE et al., 2014; GONÇALVES et al., 2009), alguns aspectos referentes a esse comércio não foram importantes aqui. A utilização de pasto em comum com outras propriedades também foi relacionada com a ocorrência da doença. Segundo Cowie et al. (2014), o compartilhamento de pastagens pode aumentar as chances de infecção do rebanho. Isso, em razão do aumento da probabilidade de contato entre rebanhos diferentes e conseqüentemente aumento das chances de introdução da doença em propriedades indenas. 1 Modelo sem pré-seleção; 2 Modelo qui-quadrado com 21 variáveis; 3 Modelo qui-quadrado com 13 variáveis

Outros fatores que apresentaram menor importância para o presente estudo, também foram verificados em estudo anterior no estado. Gonçalves et al. (2009) não encontraram associação entre a presença de brucelose em Minas Gerais com características produtivas do rebanho, como, aptidão para leite ou corte e raça dos bovinos.

Tabela 8 - Comparação entre as principais variáveis qualitativas selecionadas pelo Teste de Qui-quadrado de Pearson e Discriminante de Fisher (Continua)

Variáveis	Teste de Qui-quadrado de Pearson		Teste Discriminante de Fisher
	Valor p	Odds Ratio	Valores de J
Piquete de Parição	0,000	2,324 (1,489-3,627)	0,0901
Teste de Brucelose	0,000	2,542 (1,534-4,212)	0,0658
Tipo de assistência veterinária	0,775	---	0,0503**
Vacina Brucelose dicotômico	0,027	6,355 (0,878-45,979)	0,0487
Assistência Veterinária	0,001	2,067(1,315-3,247)	0,0448
VR leilões	0,001	3,874 (1,598-9,392)	0,0321
CR	0,028	1,632 (1,049-2,540)	0,0304
VR	0,022	1,728 (1,078- 2,771)	0,0298
Inseminação dicotômico	0,019	1,889 (1,103-3,236)	0,0294
Região	0,372	---	0,0292**
Local de abate de reprodutores	0,001	---	0,0276
Presença de Ovinos/caprinos	0,018	2,236 (1,126- 4,440)	0,0246
Resfria leite	0,064	1,513(0,973-2,354)	0,0213
Destino aborto	0,067	---	0,0212
Finalidade queijo manteiga	0,634	---	0,0202**
Tipo de ordenha	0,017	---	0,0192
Fabrica queijo/ manteiga	0,092	0,641(0,381-1,080)	0,0189
Raças bovinas	0,056	---	0,0175
Presença de gato	0,101	---	0,0172**
Consome leite cru	0,049	1,563(0,999-2,445)	0,0163*
Aluga pasto	0,095	1,496 (0,930-2,408)	0,0161*
VR fazendas	0,090	1,585 (0,927 - 2,709)	0,0157*
Regularidade teste brucelose	>0,20	---	0,0153**
Criação	>0,20	---	0,0150**
CR leilão	0,058	2,418 (0,941- 6,211)	0,0145
Destino leite	0,096	---	0,0144
Como resfria leite	>0,20	---	0,0140**
CR Fazenda	0,198	---	0,0100**
CR Exposição	0,096	---	0,0070**

*Variáveis selecionadas entre as 21 mais relevantes no Teste Qui-quadrado, porém não selecionadas pelo Fisher; ** Variáveis selecionadas entre as 30 mais relevantes no Teste de Fisher, porém não selecionadas pelo Qui-quadrado.

Legenda: CR = compra reprodutor; VR= vende reprodutor

5.5 Redes Neurais Artificiais: arquitetura e desempenho

Os melhores modelos de Redes NA foram selecionados, por meio de testes com diferentes números de neurônios na camada intermediária, selecionando aquelas com maiores níveis de sensibilidade e especificidade. Segundo Cunha et al. (2010), a maior dificuldade da Rede NA é obter os parâmetros que melhor se comportam com os dados apresentados. Parâmetros, tais como: taxas de aprendizagem, momentum, meta, quantidade de neurônios na camada escondida, quantidade de ciclos e o algoritmo de treinamento; foram obtidos após exaustivos testes.

As Redes NA têm uma capacidade de aproximação universal e podem aproximar quase qualquer tipo de função não-linear (DESAI et al., 2008; GULATI et al., 2010). No entanto, uma limitação de sua técnica de modelagem é que o tipo de sua abordagem é como um sistema de “caixa preta”, que não oferece uma visão sobre o funcionamento interno do modelo ou informação para avaliar a interação de entradas (DAYHOFF; DELEO, 2001). Dessa forma, foi de grande importância, neste trabalho, o uso da análise univariada e do discriminante de Fisher anterior ao desenvolvimento do sistema neural, como forma de maior conhecimento sobre as relações entre as variáveis e o status dos rebanhos com relação a brucelose.

As Redes NA, com ou sem pré-seleção de variáveis, mostraram-se satisfatórias, o que pode ser visto pelos níveis de sensibilidade e especificidade apresentados. A sensibilidade, no presente trabalho, foi de 83,33 para todos os modelos testados, a especificidade variou entre 53,89 a 66,60% e a eficiência de 68,61 a 74,96% (Tabela 9). O critério de seleção de uma rede pode ser obtido de modo pragmático, ou seja, a que melhor atinge os resultados esperados (CUNHA et al., 2010). E em problemas de classificação binária o seu desempenho usualmente é medido por meio da taxa de classificação correta para a amostra de teste e também calculando-se algumas estatísticas descritivas, tais

como falso-positivos, falso-negativos, sensibilidade e especificidade, que são medidas bastante utilizadas na área médica (SANTOS et al., 2003).

Um método de grande utilidade para tornar a arquitetura da rede mais prática e simples é a identificação das variáveis mais relevantes para a rede. Neste estudo, não houve diferença significativa entre os modelos, com relação à sensibilidade e eficiência, e ainda o modelo com 10 variáveis apresentou maior especificidade em relação a outros, o que permite a exclusão de parâmetros menos relevantes (Figura 7). Essa exclusão pode auxiliar na aplicabilidade e favorecer o uso das redes como um método mais simples e direto. Stark e Pfeiffer (1999) também utilizaram análise de sensibilidade para identificar as variáveis influentes, excluindo aquelas que exerciam menor influência sobre os resultados.

Não houve diferença significativa quanto à sensibilidade e eficiência entre os modelos (IC de 95%) (tabela 9 e Figura 7). O maior valor de especificidade apresentado foi o do modelo 2, que é feito por pré-seleção pelo teste qui-quadrado com 90% de confiança (Tabela 9). Porém, apenas o modelo 4 apresentou diferença significativa na média das especificidades de todos os outros modelos, exceto o 3 (figura 7). Dessa forma, demonstrou-se que os vários pré-processamentos dos dados não resultaram em diferenças importantes nos modelos. Dessa forma, considera-se que os modelos 3 e 6, que são mais simples (menor número de variáveis) devem ser preferidos.

Tabela 9 - Comparação entre a sensibilidade e especificidade das Redes Neurais Artificiais desenvolvidas por meio dos diferentes métodos de pré-seleção.

Método de pré-seleção	Modelo	Nº variáveis	Sensibilidade (%)¹	Especificidade (%)²	Eficiência (%)¹
Sem pré-seleção	1	49	83,33	61,51	72,42
Teste Qui-quadrado de Pearson*	2	21	83,33	66,60	74,96
	3	13	83,33	59,53	71,43
	4	30	83,33	61,51	72,42
Discriminante Linear de Fisher**	5	20	83,33	53,89	68,61
	6	10	83,33	61,92	72,62

*Variáveis selecionadas por meio do Teste Qui-quadrado $p < 0,1$ (seleção de 21 variáveis) e $p < 0,05$ (seleção de 13 variáveis); ** Variáveis selecionadas por meio do valor de J, onde J representa o poder de discriminação entre classes.

Equações: Sensibilidade= $VP / (VP + FN)$, VP =Verdadeiros positivos; FN=Falsos Negativos;

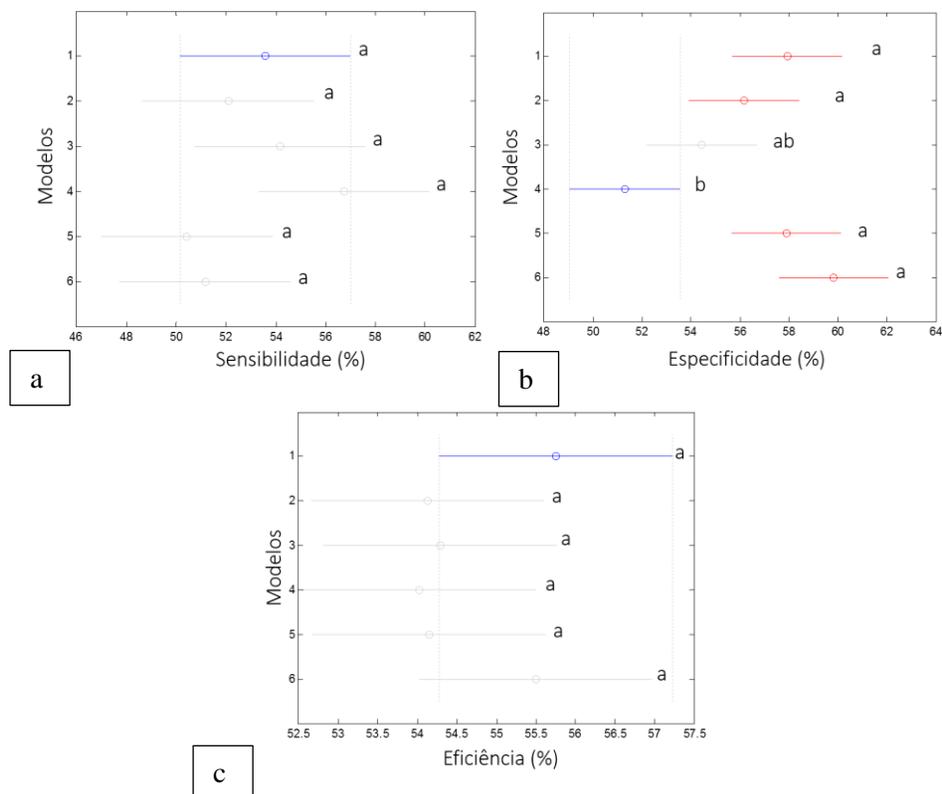
Especificidade= $VN / (VN + FP)$, VN =Verdadeiros negativos; FP=Falsos Positivos;

Eficiência= média dos níveis de sensibilidade e especificidade.

¹Não há diferença significativa ($p > 0,05$); ²- diferença significativa apenas no modelo 4 ($p < 0,05$) com todos, exceto o modelo 3.

Na Figura 7, demonstra-se a sensibilidade, especificidade e eficiência dos modelos com seus desvios-padrão. Isso facilita a demonstração da falta de diferença significativa entre a maioria.

Figura 7 - Gráficos de comparação dos valores médios e Intervalos de Confiança (IC 95%) entre os diferentes modelos de pré-processamento utilizados para o desenvolvimento das Redes Neurais Artificiais, Teste de Tukey.



Legenda: a) Gráfico de comparação da sensibilidade (IC 95%); b) Gráfico de comparação da especificidade (IC 95%); c) Gráfico de comparação da eficiência* (IC 95%); *Eficiência= média dos níveis de sensibilidade e especificidade. Diferentes letras representam diferenças significativas de média;

1 Modelo sem pré-seleção; 2 Modelo qui-quadrado com 21 variáveis; 3 Modelo qui-quadrado com 13 variáveis; 4 Modelo Fisher 30 variáveis; 5 Modelo Fisher 20 variáveis; 6 Modelo Fisher 10 variáveis;

Os modelos de Rede NA exigem mais experiências (número de padrões) para construir um modelo eficiente que outros testes estatísticos. Assim, o número reduzido de observações do grupo positivo para brucelose e, ainda, o grande desequilíbrio do número de observações entre o grupo positivo e

negativo trouxeram limitações aos resultados no presente estudo. Apesar de alguns autores citarem que o modelo de Rede NA pode funcionar bem, com um número menor de dados, se estes forem bem distribuídos no domínio de entrada (DESAI et al., 2008; LOU; NAKAI, 2001), seria interessante testar esses modelos em bancos de dados maiores e em doenças que apresentem diferenças mais marcantes entre rebanhos positivos e negativos.

No presente estudo, optou-se pela replicação, ou sobreamostragem, dos dados do conjunto positivo e a subamostragem aleatória do conjunto negativo. O intuito desses recursos foi justamente minimizar o desequilíbrio entre as classes e facilitar o aprendizado da rede. Porém, a subamostragem gera múltiplas cópias do mesmo padrão e isso traz algumas desvantagens, como dificuldade de reconhecimento da classe minoritária e aumento da variância, observados neste estudo (BATISTA, 2003; CASTRO; BRAGA, 2011).

Os desempenhos obtidos para a rede com melhor resultado evidenciou a generalização do método, em que desempenhos próximos aos do conjunto de dados do projeto foram alcançados, levando a uma média de acertos totais entre 68,61 a 74,96% como observado na tabela 9. Tanto os resultados de desempenho para os dados do projeto, quanto para os dados de validação apresentaram maior variação de uma execução para outra no grupo positivo, quando comparados ao grupo negativo, conforme demonstrado pelos desvios padrão, que variaram de 12,53 a 18,37% e 8,62 a 15,16% para os dois grupos, respectivamente. Os desvios padrão maiores obtidos para o grupo positivo podem ser justificados pelo reduzido número de casos dessa classe, comparado com a classe de propriedades negativas. Por outro lado, desvios padrão altos podem indicar maior dispersão da classe no espaço de parâmetros, ou seja, são geralmente encontrados para classes mais heterogêneas.

A maior variação dos dados pode comprometer a repetibilidade da técnica e isso pode ser decorrente do fato do treinamento ter sido realizado com

poucos dados e ainda pelo fato da utilização da técnica de replicação de dados, no caso do grupo soropositivo. Ghotoorlar et al. (2012), trabalhando com um grupo de 105 vacas leiteiras e 23 características, encontraram valores de sensibilidade e especificidade maiores do que 72% em Rede NA para detecção de laminite em rebanhos leiteiros. Tais resultados assemelham-se ao presente estudo, porém, o desvio padrão médio foi de 0,31 e o coeficiente médio de variação foi de 14,55%, o que representou uma repetibilidade elevada, que se aproxima aos resultados do grupo negativo desse trabalho.

O número de observações disponíveis para análise, realmente pode ser um fator limitante para os modelos de Redes NA, porém as características desses dados, tais como, seu poder de informação e sua variabilidade também têm grande influência no aprendizado de uma Rede NA. Alguns trabalhos, utilizando um número de observações semelhantes ao presente estudo, obtiveram resultados satisfatórios, utilizando as Redes NA no auxílio de classificações, diagnósticos e previsões, na área veterinária (ANKINAKATTE et al. 2013; CARSON et al., 1995; GHOTOORLAR et al., 2012; MCEVOY; AMIGO, 2013; ROUSH et al., 1997).

Algumas Redes NA apresentaram-se muito eficientes na literatura, alcançando níveis altos de sensibilidade e especificidade. Como exemplo, o uso das Redes NA para reconhecimento da estirpe de *Escherichia coli* O157: H7, pareceu promissor. O valor preditivo positivo foi de 91,6%, e o valor preditivo negativo foi de 96,2% , sensibilidade 97,0% e a especificidade 90,9% (CARSON et al., 1995). Roush et al. (1997) utilizaram Redes NA para a escolha de frangos com ausência de propensão para a ascite também foi eficiente. Os autores encontraram uma sensibilidade de 95% (19/20 aves) e especificidade de 88% (14/16 aves) e, para isso, utilizaram 170 observações (treinamento 117 aves; calibragem 17 aves; verificação 36 aves) (ROUSH et al., 1997). Para a classificação de radiografias caninas, foram utilizadas 200 imagens para treino

da rede (120 quadril e 80 não quadril) e 56 imagens de teste (36 de quadril e 20 não quadril). A Rede NA apresentou o erro de classificação de 8,9%, a sensibilidade de 86%, e a especificidade de 100%. Segundo os autores, as redes neurais têm o potencial para agrupamento e classificação, mas é importante que se disponha de um grande número de dados de boa qualidade para treinamento do modelo (MCEVOY; AMIGO, 2013).

Ankinakatte et al. (2013), com o objetivo de uma detecção precoce de mastite em vacas, utilizaram dados de 87 lactações, e encontraram uma sensibilidade de quase 75%, para 80% de especificidade fixa. Contudo, Cavero et al. (2008), com o mesmo propósito, utilizaram um amplo conjunto de dados, envolvendo 403.537 ordenhas e 478 vacas, mas o desempenho do Sistema Neural não foi satisfatório. O número médio de verdadeiros positivos por dia variou entre 1,2-6,4, e a média do número de falsos positivos variou entre 5,2-6,8, em um rebanho de 24 vacas por dia. Segundo os autores, uma diminuição da taxa de erro poderia ser conseguida por meio de parâmetros mais informativos. Os resultados apresentados, nestes estudos, ratificam a importância, não só do número de observações disponíveis, como também da qualidade das informações fornecidas, para um bom desempenho do Sistema Neural.

O modelo de rede proposto pode ser implementado em computadores básicos, utilizando o próprio software MatLab ou *softwares* mais populares, como o Excel, já que o processamento matemático envolvido é bastante simples, em razão da estrutura simples do modelo, que envolve operações matemáticas básicas como adição, subtração, multiplicação e tangente hiperbólica. Para exemplificação, a Rede NA, baseada no Discriminante de Fisher com dez variáveis, foi implementada em *software* Excel, por meio do qual foi possível inserir os dados referentes a cada propriedade para sua classificação quanto à doença.

Assim, testou-se sua aplicabilidade para a segregação das propriedades quanto ao *status* da doença (Figura 8). A opção pelo modelo composto por dez variáveis foi pelo fato da maior praticidade e aplicabilidade em trabalhar com um menor número de variáveis, já que estas foram suficientes para a eficiência da rede, e ainda foi o modelo com um dos maiores níveis de especificidade. Os resultados disponibilizados pela rede podem ser tanto qualitativos, quanto quantitativos, como exemplificado na Figura 8. Assim, quando o rebanho for classificado como positivo, a saída da rede será um valor positivo, e quando o rebanho enquadrar-se como negativo pela rede a saída será um valor negativo.

Além disso, a rede proporciona uma saída numérica, em uma escala que vai de -1 a 1, ou seja, quanto mais próximo de -1 mais as características da propriedade se enquadram no padrão negativo estipulado pela rede, e quanto mais próxima de 1 mais a propriedade engloba características de um padrão positivo para a doença. Dessa forma, há um grau da certeza da informação fornecida pelo modelo neural, no que tange à classificação de risco da propriedade processada (figura 8).

As variáveis mais relevantes para o modelo, relacionadas às características da propriedade (número de animais, piquete de parição), e manejo do rebanho (teste de brucelose, assistência veterinária, vacinação contra brucelose, comércio de animais, etc.) poderiam ser coletadas facilmente em uma busca ativa ou, até melhor, por busca passiva, se as fichas cadastro dos órgãos de defesa contemplassem esses tipos de informações dos produtores rurais. Dessa forma, demonstra-se que a Rede NA pode ser um meio interessante para a triagem de propriedades de maneira mais econômica e viável, permitindo que os órgãos de fiscalização concentrem suas ações naquelas propriedades com maior probabilidade de possuir animais soropositivos para brucelose.

6 CONSIDERAÇÕES FINAIS

Como projeções futuras, consideramos que, com o uso das Redes NA, o nível de risco do rebanho pode ser enquadrado em dois grupos: negativo (baixo risco) e positivo (alto risco) para brucelose bovina. Informações estatísticas acerca da posição em que o rebanho se encontra em cada um desses grupos também podem ser disponibilizadas como resultados numéricos e/ou gráficos.

Por ser um trabalho pioneiro na área da defesa sanitária animal, o conhecimento adquirido no desenvolvimento e adequação do sistema poderá ser aplicado para o diagnóstico de outras doenças relevantes na medicina veterinária, principalmente para aquelas que têm prevalência maior, o que, provavelmente, traz melhoria à eficiência do modelo, por um maior equilíbrio entre positivos e negativos.

7 CONCLUSÃO

Os resultados demonstram uma condição prática de aplicação das Redes NA em discriminação de rebanhos positivos e negativos na defesa animal para predição de rebanhos positivos ou negativos. Essa aplicação para brucelose parece promissora, com eficiência acima de 70%.

Os parâmetros que mais contribuíram para a discriminação dos positivos e negativos nos modelos de redes neurais para brucelose bovina foram aqueles relacionados ao tamanho do rebanho, ao manejo sanitário e ao comércio de animais. Estas são as informações mais relevantes para na triagem de rebanhos positivos.

Não foi encontrada diferença significativa quanto aos níveis de sensibilidade e eficiência entre as diferentes pré-seleções utilizadas. Quanto à especificidade, não houve diferença relevante entre a maioria dos modelos. A comparação entre estes foi interessante por permitir a escolha de um modelo mais simples (com menor número de variáveis), considerando a manutenção da eficiência.

REFERÊNCIAS

- ACHA, P. N.; SZYFRES, B. **Zoonosis y enfermedades transmisibles comunes al hombre y a los animales: bacterioses e micoses**. 3. ed. Washington: OPAS, 2001. 416 p.
- AHMADI, H.; GOLIAN, A. Response surface and neural network models for performance of broiler chicks fed diets varying in digestible protein and critical amino acids from 11 to 17 days of age. **Poultry Science**, Champaign, v. 90, n. 9, p. 2085–2096, Sept. 2011.
- ALVES, A. J. S et al. Economic analysis of vaccination to control bovine brucellosis in the States of Sao Paulo and Mato Grosso, Brazil. **Preventive Veterinary Medicine**, Amsterdam, v. 118, n. 4, p. 351–358, Mar. 2015.
- ANKINAKATTE, S. et al. Predicting mastitis in dairy cows using neural networks and generalized additive models: a comparison. **Computers and Electronics in Agriculture**, New York, v. 99 p. 1–6, Nov. 2013.
- ARENAS, G. N. et al. Intracellular trafficking of *Brucella abortus* J774 macrophages. **Infection and Immunity**, Washington, v. 68, n. 7, p. 4255–4263, July 2000.
- ASMARE, K. et al. The status of bovine brucellosis in Ethiopia with special emphasis on exotic and cross bred cattle in dairy and breeding farms. **Acta Tropica**, Basel, v. 126, n. 3, p. 186–192, June 2013.
- AZEVEDO, S. S. et al. Situação epidemiológica da brucelose bovina no Estado do Espírito Santo. **Arquivo Brasileiro de Medicina Veterinária e Zootecnia**, Belo Horizonte, v. 61, supl. 1, p. 19-26, nov. 2009.
- BALDWIN, C. L.; GOENKA, R. Host immune responses to the intracellular bacteria *Brucella*: does the bacteria instruct the host to facilitate chronic infection? **Critical Reviews in Immunology**, Boca Raton, v. 26, n. 5, p. 407–442, 2006.
- BALDWIN, C. L.; WINTER, A. J. Macrophages and *Brucella*. **Immunology Series**, New York, v. 60, p. 363–380, 1994.
- BATISTA, G. **Pré-processamento de dados em aprendizado de máquina supervisionado**. São Carlos: ICMC, 2003. 206 p.

BERCOVICH, Z. Maintenance of *Brucella abortus*-free herds: a review with emphasis on the epidemiology and the problems in diagnosing brucellosis in areas of low prevalence. **The Veterinary Quarterly**, The Hague, v. 20, n. 3, p. 81-88, July 1998.

BORBA, M. R. et al. Prevalence and risk-mapping of bovine brucellosis in Maranhão State, Brazil. **Preventive Veterinary Medicine**, Amsterdam, v. 110, n. 2, p. 169–176, June 2013.

BRAGA, A. P.; CARVALHO, A. C. P. L. F.; LUDERMIR, T. B. **Redes neurais artificiais: teoria e aplicações**. 2. ed. Rio de Janeiro: LTC, 2007. 238 p.

BRASIL. Ministério da Agricultura e Abastecimento. **Diagnóstico de saúde animal**. Brasília: MAPA, 1977. 735 p.

_____. Instrução Normativa nº 6, de 8 de janeiro de 2004. Aprova o Regulamento Técnico do Programa Nacional de Controle e Erradicação da Brucelose e Tuberculose Animal (PNCEBT). **Diário Oficial da União**, Brasília, DF, n. 7, p. 6-10, jan. 2004. Seção 1.

_____. Departamento de Defesa Animal. **Regulamento técnico do programa nacional de controle e erradicação da brucelose e tuberculose – PNCEBT**. Brasília: MAPA, 2001. 21 p.

_____. **Manual de procedimentos do estudo epidemiológico sobre brucelose e tuberculose em bovinos e bubalinos**. Brasília: MAPA, 2009. 188 p.

_____. Programa Nacional de Controle e Erradicação da Brucelose e Tuberculose. **Manual técnico**. Brasília : MAPA, 2006. 184 p.

CALISTRI, P. et al. Risk factors for the persistence of bovine brucellosis in Sicily from 2008 to 2010. **Preventive Veterinary Medicine**, Amsterdam, v. 110, n. 3/4, p. 329–334, July 2013.

CARSON, C. A. *Escherichia coli* O157:H7 restriction pattern recognition by artificial neural network. **Journal of Clinical Microbiology**, Washington, v. 33, n. 11, p. 2894–2898, Nov. 1995.

CASTRO, C. L.; BRAGA, A. P. Aprendizado supervisionado com conjuntos de dados desbalanceados. **SBA Controle & Automação**, Campinas, v. 22, n. 5, p. 441-466, set./out. 2011.

CAVERO, D. et al. Mastitis detection in dairy cows by application of neural networks. **Livestock Science**, Amsterdam, v. 114, n. 2/3, p. 280-286, Apr. 2008.

CHATE, S. C. et al. Situação epidemiológica da brucelose bovina no Estado do Mato Grosso do Sul. **Arquivo Brasileiro de Medicina Veterinária e Zootecnia**, Belo Horizonte, v. 61, supl. 1, p. 46-55, nov. 2009.

CORBEL, M. J. Brucellosis: an overview. **Infectious Diseases Journals**, Atlanta, v. 3, n. 2, p. 213-221, Apr./June 1997.

COWIE, C. E. A. et al. Shared risk factors for multiple livestock diseases: a case study of bovine tuberculosis and brucellosis. **Research in Veterinary Science**, Oxford, v. 97, n. 3, p. 491-497, Dec. 2014.

CUNHA, G. B. et al. A utilização de uma rede neural artificial para previsão da incidência da malária no Município de Canta, Estado de Roraima. **Revista da Sociedade Brasileira de Medicina Tropical**, Uberaba, v. 43, n. 5, p. 567-570, set./out. 2010.

D'ANASTASIO, R. et al. Origin, evolution and paleoepidemiology of brucellosis. **Epidemiology and Infection**, Cambridge, v. 139, n. 1, p. 149-156, Jan. 2011.

DAYHOFF, J. E.; DELEO, J. M. Artificial neural networks: opening the black box. **Cancer**, New York, v. 91, n. 8, p. 1615-1635, Apr. 2001.

DESAI, K. M. et al. Comparison of artificial neural network (ANN) and response surface methodology (RSM) in fermentation media optimization: case study of fermentative production of scleroglucan. **Biochemical Engineering Journal**, Amsterdam, v. 41, n. 3, p. 266-273, Oct. 2008.

DORNELES, E. M. S. et al. Immune response triggered by *Brucella abortus* following infection or vaccination. **Vaccine**, Guildford, v. 33, n. 31, p. 3659-3666, July 2015.

DUDA, R. O.; HART, P. E.; STORK, D. G. **Pattern classification**. 2. ed. New York: Wiley-Interscience, 2000. 680 p.

FRETIN, D. et al. The sheathed flagellum of *Brucella melitensis* is involved in persistence in a murine model of infection. **Cellular Microbiology**, Oxford, v. 7, n. 5, p. 687-698, May 2005.

GAITA, L.; MILITARU, M. Pathology and clinical data integrated in artificial neural networks for providing prognosis in câncer. **Journal of Comparative Pathology**, Edinburgh, v. 148, n. 1, p. 214-219, Jan. 2013.

GHOTOORLAR, S. M. et al. Lameness scoring system for dairy cows using force plates and artificial intelligence. **Veterinary Record**, London, v. 170, n. 5, p. 126-130, Feb. 2012.

GODFROID, J. et al. How to substantiate eradication of bovine brucellosis when specific serological reactions occur in the course of brucellosis testing. **Veterinary Microbiology**, Amsterdam, v. 90, n. 1/4, p. 461-477, Dec. 2002.

GONÇALVES, V. S. P. et al. Situação epidemiológica da brucelose bovina no Estado de Minas Gerais. **Arquivo Brasileiro de Medicina Veterinária e Zootecnia**, Belo Horizonte, v. 61, supl. 1, p. 35-45, nov. 2009.

GORVEL, J. P. Brucella: a Mr “Hide” converted into Dr Jekyll. **Microbes and Infection**, Paris, v. 10, n. 9, p. 1010-1013, July 2008.

GORVEL, J. P.; MORENO, E. *Brucella* intracellular life: from invasion to intracellular replication. **Veterinary Microbiology**, Amsterdam, v. 90, n. 1/4, p. 281-297, Dec. 2002.

GULATI, T. et al. Comparative study of response surface methodology, artificial neural network and genetic algorithms for optimization of soybean hydration. **Food Technology and Biotechnology**, Washington, v. 48, p. 11–18, June 2010.

HAYKIN, S. **Neural networks and learning machines**. 3. ed. New York: Prentice Hall, 2008. 936 p.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA - IBGE. Estatística da produção pecuária: pecuárias: produção da Pecuária Municipal – 2015. **Estados@**, Brasília, 2015. Disponível em: <[http://www.ibge.gov.br/estadosat/temas.php?sigla=mg&tema=pecuaria 2015](http://www.ibge.gov.br/estadosat/temas.php?sigla=mg&tema=pecuaria%2015)>. Acesso em: 11 out. 2016.

_____. **Estatística da produção pecuária: censo agropecuário 2006: Brasil, grandes regiões e unidades da federação: segunda apuração**. Rio de Janeiro: IBGE, 2012. 774 p.

_____. **Indicadores IBGE**: setembro de 2016. Brasília: IBGE, 2016. 45 p.

JANE, Y. N.; NEHEMIAH, H. K.; ARPUTHARAJ, K. A Q-backpropagated time delay neural network for diagnosing severity of gait disturbances in Parkinson's disease. **Journal of Biomedical Informatics**, San Diego, v. 60, p. 169–176, Apr. 2016.

KÖHLER, S. et al. The analysis of the intramacrophagic virulome of *Brucella suis* deciphers the environment encountered by the pathogen inside the macrophage host cell. **Proceedings of the National Academy of Sciences of the United States of America**, Washington, v. 99, n. 24, p. 15711–15716, Nov. 2002.

_____. What is the nature of the replicative niche of a stealthy bug named *Brucella*? **Trends in Microbiology**, Cambridge, v. 11, n. 5, p. 215–219, May 2003.

LOU, W.; NAKAI, S. Application of artificial neural networks for predicting the thermal inactivation of bacteria: a combined effect of temperature, pH and water activity. **Food Research International**, Barking, v. 34, n. 7, p. 573-579, 2001.

MAIDANTCHIK, C. L. L. et al. Neuraltb web system support to the smear negative pulmonary tuberculosis diagnosis. In: INTERNATIONAL CONFERENCE ON ENTERPRISE INFORMATION SYSTEMS, 9., 2007, Funchal. **Proceedings...** Funchal: [s.n.], 2007. p. 198-203.

MCEVOY, F. J.; AMIGO, J. M. Using machine learning to classify image features from canine pelvic radiographs: evaluation of partial least squares discriminant analysis and artificial neural network models. **Veterinary Radiology & Ultrasound**, Raleigh, v. 54, n. 2, p. 122–126, 2013.

MEDRONHO, R. A. et al. **Epidemiologia**. 2. ed. São Paulo: Atheneu, 2009. 790 p.

MEGID, J. et al. Avaliação das provas de soroaglutinação rápida, soro aglutinação lenta, antígeno acidificado e 2-mercaptoetanol no diagnóstico da brucelose bovina. **Brazilian Journal of Veterinary Research and Animal Science**, São Paulo, v. 37, n. 5, p. 395-399, 2000.

MEMISH, Z. A.; BALKHY, H. H. Brucellosis and international travel. **Journal of Travel Medicine**, Hamilton, v. 11, n. 1, p. 49-55, Jan./Feb. 2004.

MOLLER, M. F. A scaled conjugate gradient algorithm for fast supervised learning. **Neural Networks**, New York, v. 6, n. 4, p. 525-533, 1993.

NAROENI, A. et al. *Brucella suis* - impaired specific recognition of phagosomes by lysosomes due to phagosomal membrane modifications. **Infection and Immunity**, Washington, v. 69, n. 1, p. 486-493, Jan. 2001.

NEGREIROS, R. L. et al. Situação epidemiológica da brucelose bovina no Estado de Mato Grosso. **Arquivo Brasileiro de Medicina Veterinária e Zootecnia**, Belo Horizonte, v. 61, supl. 1, p. 56-65, Nov. 2009.

NIELSEN, K. Diagnosis of brucellosis by serology. **Veterinary Microbiology**, v. 90, n. 1/4, p. 447-459, Dec. 2002.

NIELSEN, K.; GALL, D. Fluorescence polarization assay for the diagnosis of brucellosis: a review. **Journal of Immunoassay & Immunochemistry**, New York, v. 22, n. 3, p. 183-201, 2001.

NOORDHUIZEN, J. P. T. M. et al. **Application of quantitative methods in veterinary epidemiology**. Wageningen: Wageningen Press, 1997. 445 p.

OCHOLI, R. A. et al. Phenotypic characterization of *Brucella* strains isolated from livestock in Nigeria. **Veterinary Microbiology**, Amsterdam, v. 103, n. 1/2, p. 47-53, Oct. 2004.

OLIVEIRA, L. F. et al. Seroprevalence and risk factors for bovine brucellosis In Minas Gerais State, Brazil. **Semina: ciências agrárias**, Londrina, v. 37, supl. 2, p. 3449-3466, 2016.

PAPPAS, G. The changing *Brucella* ecology: novel reservoirs, new threats. **International Journal of Antimicrobial Agents**, Amsterdam, v. 36, Suppl. 1, p. 8-11, Nov. 2010.

PAULIN, L. M. et al. Estudo comparativo dos testes de 2- mercaptoetanol e reação de fixação do complemento no sorodiagnóstico da brucelose bovina. **Arquivos do Instituto Biológico**, São Paulo, v. 69, n. 4, p. 41- 47, out./dez. 2002.

PAULIN, L. M.; FERREIRA NETO, J. S. **O combate à brucelose bovina: situação brasileira**. Jaboticabal: Funep, 2003. 154 p.

PEREIRA, M. G. **Epidemiologia: teoria e prática**. Rio de Janeiro: Guanabara Koogan, 2005. 598 p.

POESTER, F. P. et al. Estudos de prevalência da brucelose bovina no âmbito do Programa Nacional de Controle e Erradicação de Brucelose e Tuberculose: introdução. **Arquivo Brasileiro de Medicina Veterinária e Zootecnia**, Belo Horizonte, v. 61, supl. 1, p. 1-5, nov. 2009.

POESTER, F. P.; GONÇALVES, V. S. P.; LAGE, A. P. Brucellosis in Brazil. **Veterinary Microbiology**, Amsterdam, v. 90, n. 1/4, p. 55-62, Dec. 2002.

PRECHELT, L. Automatic early stopping using cross validation: quantifying the criteria. **Neural Networks**, New York, v. 11, n. 4, p. 761-767, June 1998.

ROCHA, W. V. et al. Situação epidemiológica da brucelose bovina no Estado de Goiás. **Arquivo Brasileiro de Medicina Veterinária e Zootecnia**, Belo Horizonte, v. 61, supl. 1, p. 27-34, nov. 2009.

ROSENBLATT, F. **Principles of neurodynamics**: perceptrons and the theory of brain mechanisms. Michigan: Spartan Books, 1962. 616 p.

ROTHMAN, K. J.; GREENLAND, S. **Modern epidemiology**. 2. ed. Philadelphia: Lippincot Willians & Wilkins, 1998. 738 p.

ROUQUAYROL, M. Z.; GURGEL, M. **Epidemiologia & saúde**. 7. ed. Rio de Janeiro: Medbook, 2013. 736 p.

ROUSH, W. B. et al. Artificial neural network prediction of ascites in broilers. **Poultry Science**, Champaign, v. 75, n. 12, p. 1479-1487, Dec. 1996.

_____. Probabilistic neural network prediction of ascites in broilers based on minimally invasive physiological factors. **Poultry Science**, Champaign, v. 76, p. 1513-1516, 1997.

SANOGO, M. et al. **Risk factors associated with brucellosis seropositivity among cattle in the central savannah-forest area of Ivory Coast**. *Preventive Veterinary Medicine*, Amsterdam, v. 107, n. 1/2, p. 51-56, Nov. 2012.

SANTOS, A. M. et al. Árvore de classificação e redes neurais artificiais: uma aplicação à predição de tuberculose pulmonar. In: CONGRESSO BRASILEIRO DE REDES NEURAIAS, 6., 2003, São Paulo. **Anais...** São Paulo: CBRN, 2003. p. 427-432.

_____. Usando redes neurais artificiais e regressão logística na predição da hepatite A. **Revista Brasileira de Epidemiologia**, São Paulo, v. 8, n. 2, p. 117-126, 2005.

_____. Economic losses due to bovine brucellosis in Brazil. **Pesquisa Veterinária Brasileira**, Rio de Janeiro, v. 33, n. 6, p. 759-764, jun. 2013.

SCHLABRITZ-LOUTSEVITCH, N. E. et al. A novel *Brucella* isolate in association with two cases of stillbirth in non-human primates: first report. **Journal of Medical Primatology**, Basel, v. 38, n. 1, p. 70–73, Feb. 2009.

SCHOBESBERGER, H.; PEHAM, C. Computerized detection of supporting forelimb lameness in the horse using an artificial neural network. **The Veterinary Journal**, London, v. 163, n. 1, p. 77-84, Jan. 2002.

SCHOLZ, H. C. et al. *Brucella inopinata* sp. nov., isolated from a breast implant infection. **International Journal of Systematic and Evolutionary Microbiology**, Reading, v. 60, n. 4, p. 801-808, Abr. 2010.

_____. Isolation of *Brucella microti* from soil. **Emerging Infectious Diseases Journal**, Atlanta, v. 14, n. 8, p. 1316–1317, Aug. 2008.

SEIXAS, J. M. et al. Neural Networks: an application for predicting smear negative pulmonary tuberculosis. In: BALAKRISHNAN, N. et al. (Org.). **Advances in statistical methods for the health sciences**. Alemanha: Hardcover, 2006. p. 279-292.

SIKUSAWA, S. et al. Situação epidemiológica da brucelose bovina no Estado de Santa Catarina. **Arquivo Brasileiro de Medicina Veterinária e Zootecnia**, Belo Horizonte, v. 61, supl. 1, p. 103-108, nov. 2009.

SILVA JÚNIOR, F. F. et al. Avaliação do teste do anel em leite na vigilância epidemiológica da brucelose bovina em rebanhos e em laticínios. **Arquivo Brasileiro de Medicina Veterinária e Zootecnia**, Belo Horizonte, v. 59, n. 2, p. 295-300, abr. 2007.

STÄRK, K. D. C.; PFEIFFER, D. U. The application of non-parametric techniques to solve classification problems in complex data sets in veterinary epidemiology - an example. **Intelligent Data Analysis**, New York, v. 3, n. 1, p. 23-35, May 1999.

THE WORLD ORGANISATION FOR ANIMAL HEALTH. Bovine brucellosis. **Manual of diagnostic tests and vaccines for terrestrial animals**. New York: OIE, 2012. p. 616-650.

_____. **Manual of standards for diagnostic tests and vaccines**. 3. ed. Paris: Office International Des Epizooties, 1996. 1343 p.

THEODORIDIS, S.; KOUTROUMBAS, K. **Pattern recognition**. 4. ed. New York: Academic Press, 2009. 984 p.

TILLER, R. V. et al. Identification of an unusual *Brucella strain* (BO2) from a lung biopsy in a 52 year-old patient with chronic destructive pneumonia. **BMC Microbiology**, London, v. 10, p. 1-11, Jan. 2010.

VERGER, J. M. et al. Taxonomy of the genus *Brucella*. **Annales de l'Institut Pasteur Microbiology**, Paris, v. 138, n. 2, p. 235-238, Mar./Apr. 1987.

VERONEZI, R. **Doenças infecciosas e parasitárias**. 6. ed. Rio de Janeiro: Guanabara-Koogan, 1976. 421 p.

VILLAR, K. S. et al. Situação epidemiológica da brucelose bovina no Estado de Rondônia. **Arquivo Brasileiro de Medicina Veterinária e Zootecnia**, Belo Horizonte, v. 61, supl. 1, p. 85-92, nov. 2009.

YINGST, S.; HOOVER, D. L. T cell immunity to brucellosis. **Critical Reviews in Microbiology**, Boca Raton, v. 29, n. 4, p. 313-331, 2003.