

**UNIVERSIDADE FEDERAL DE MINAS GERAIS  
ESCOLA DE VETERINÁRIA**

**AVALIAÇÃO GENÉTICA DE TOUROS DA RAÇA TABAPUÃ E  
PREDIÇÃO DA LACTAÇÃO DE ANIMAIS DA RAÇA  
GIR LEITEIRO E CRUZAMENTOS COM RAÇA  
HOLANDESA POR INTERMÉDIO DE REDES  
NEURAIS ARTIFICIAIS**

**Ricardo Vieira Ventura**

Orientador: Prof. Martinho de Almeida e Silva.

Belo Horizonte,  
2010

**Ricardo Vieira Ventura**

**AVALIAÇÃO GENÉTICA DE TOUROS DA RAÇA TABAPUÃ E  
PREDIÇÃO DA LACTAÇÃO DE ANIMAIS DA RAÇA  
GIR LEITEIRO E CRUZAMENTOS COM RAÇA  
HOLANDESA POR INTERMÉDIO DE REDES  
NEURAIS ARTIFICIAIS**

**Tese apresentada à Escola de Veterinária da  
Universidade Federal de Minas Gerais, como  
requisito parcial para obtenção do grau de  
Doutor em Zootecnia**

**Área: Genética e Melhoramento Animal**

**Orientador: Prof. Martinho de Almeida e Silva**

Belo Horizonte  
Escola de Veterinária – UFMG  
2010

V657a Ventura, Ricardo Vieira, 1980-  
Avaliação genética de touros da raça Tabapuã e predição da lactação de animais da raça Gir leiteiro e cruzamentos com raça Holandesa por intermédio de redes neurais artificiais/  
Ricardo Vieira Ventura. – 2010.  
63 p. : il.

Orientador: Martinho de Almeida e Silva  
Tese (doutorado) – Universidade Federal de Minas Gerais, Escola de Veterinária  
Inclui bibliografia

1. Tabapuã (Zebu) – Genética – Teses. 2. Bovino de leite – Melhoramento genético – Teses. 3. Algoritmo genético – Teses. 4. Redes neurais (Computação) – Teses. I. Silva, Martinho de Almeida. II. Universidade Federal de Minas Gerais. Escola de Veterinária. III. Título.

CDD – 636.208 2

Tese defendida e aprovada em 24 de fevereiro de 2010, pela Comissão Examinadora constituída por:

Martinho de Almeida e Silva



André Gustavo Pereira de Andrade



Aníbal Eugênio Vercesi Filho



Nelson José Laurino Dionello



Robledo de Almeida Torres



*À minha querida esposa, pelo incentivo e compreensão em todos os momentos e ao meu filho Rafael Antônio que acaba de nascer.*

*Aos meus pais Antônio e Maria, sendo imprescindíveis para a realização desse trabalho.*

*Aos meus padrinhos Martinho e Albertina, pela amizade sincera e apoio constante.*

*Aos meus irmãos e familiares, por estarem sempre ao meu lado, transmitindo carinho e confiança.*

*À minha sogra Léo, meu cunhado Flávio e Srta. Helen por toda a paciência e ajuda em momentos tão difíceis.*



## **AGRADECIMENTOS**

A Deus, por me proporcionar saúde e disposição, oferecendo-me oportunidade de crescimento, conhecimento e amadurecimento pessoal.

Ao grande amigo e Professor Martinho de Almeida e Silva, pelo carinho e credibilidade nunca recebidos anteriormente, transmitindo seus conhecimentos com muito empenho e honestidade. Pessoa a quem serei eternamente grato, prestativo em todos os momentos, a quem considero meu “segundo” pai. À Albertina, pelo carinho e transmissão de fé.

Ao Dr. Talles Medeiros, pelo ensinamento e ajuda na área tecnológica, pelo constante incentivo e colaboração direta nas minhas tarefas, proporcionando-me tranqüilidade para execução desse trabalho.

Aos amigos e Professores Nelson José Laurino Dionello, Robledo de Almeida Torres e Aníbal Vercesi Filho, pela amizade e ensinamentos oferecidos.

Ao Professor e grande amigo Fernando Enrique Madalena, pelas inúmeras e excelentes orientações, propiciando oportunidades raras de aprendizado, crescimento pessoal e profissional.

Aos Professores André Gustavo Pereira de Andrade, Jonas Carlos Campos Pereira e Sandra Gesteira Coelho, pela disponibilidade e colaboração.

Aos Doutores Ângela Beatriz Fridrich e Bruno Dourado Valente pela ajuda em diversos momentos e riqueza de detalhes em seus ensinamentos.

Aos professores Denise Aparecida de Andrade Oliveira, Ângela Maria Quintão Lana, pela oportunidade de trabalho em conjunto, colaborando para o meu crescimento. Ao Professor Ronaldo Braga Reis, pelo voto de confiança ao me oferecer oportunidade de ingresso na área de Veterinária.

A todos do Laticínios Chapada e Tortuga, pelo apoio.

Aos “manos” Hélio Botina, Gerusa, Gal, “Santa Luciana Freitas dos momentos impossíveis”, Joana, Mitzem e Isabel pela ajuda e companheirismo.

A todos da FEPMVZ-Editora, em especial à Claudia Kafuri.

Aos funcionários Heloísa (“Santa Heloísa”), Eliane, Nilda e Luciene, que contribuíram de alguma forma para a realização desse trabalho.

A EMBRAPA – Gado de Leite e à ABCZ (Associação Brasileira de Criadores de Zebu) por terem cedido os dados dos animais da raça Gir e Tabapuã para a realização do presente estudo.

*“A mente que se abre a uma nova idéia, nunca volta ao seu tamanho original”.*

(Albert Einstein)

## SUMÁRIO

	<b>Resumo</b> .....	12
	<b>Abstract</b> .....	12
1.	<b>INTRODUÇÃO</b> .....	13
2.	<b>REVISÃO DE LITERATURA</b> .....	16
2.1	<b>Função limiar (Degrau)</b> .....	20
2.2	<b>Função Sigmóide</b> .....	20
2.3	<b>Tangente Hiperbólica</b> .....	21
2.4	<b>Resumo do Algoritmo Backpropagation</b> .....	25
2.5	<b>Algoritmo Levenberg-Marquardt</b> .....	26
2.6	<b>Algoritmo Multi-Objetivo</b> .....	27
3.	<b>REFERENCIAS BIBLIOGRAFICAS</b> .....	29
	<b>CAPITULO 1. USO DE REDES NEURAIAS ARTIFICAIAS NA PREDIÇÃO DE VALORES GENÉTICOS PARA PESO AOS 205 DIAS EM BOVINOS DA RAÇA TABAPUÃ</b> .....	32
	<b>RESUMO</b> .....	32
	<b>ABSTRACT</b> .....	32
1.	<b>INTRODUÇÃO</b> .....	33
2.	<b>MATERIAL E MÉTODOS</b> .....	35
3.	<b>RESULTADOS</b> .....	38
4.	<b>CONCLUSÕES</b> .....	42
5.	<b>REFERENCIAS BIBLIOGRÁFICAS</b> .....	42
	<b>CAPITULO 2. PREDIÇÕES DA PRODUÇÃO DE LEITE NO SEGUNDO E TERCEIRO PARTOS POR INTERMÉDIO DE INFORMAÇÕES DOS PARTOS ANTERIORES UTILIZANDO REDES NEURAIAS ARTIFICIAIS E COMPARAÇÃO COM A CAPACIDADE PROVÁVEL DE PRODUÇÃO</b> .....	44
	<b>RESUMO</b> .....	44
	<b>ABSTRACT</b> .....	45
1.	<b>INTRODUÇÃO</b> .....	45
2.	<b>MATERIAL E MÉTODOS</b> .....	47
2.1	<b>Predição da segunda lactação por meio de informações da primeira lactação</b> .....	48
2.2	<b>Predição da terceira lactação por meio de informações da primeira e segunda lactação</b> .....	51
3	<b>RESULTADOS</b> .....	52
3.1	<b>Resultados - Predição da segunda lactação por meio de informações da primeira lactação</b> .....	52
3.2	<b>Resultados - Predição da terceira lactação por meio de informações da primeira e segunda lactação</b> .....	58
4	<b>CONCLUSÕES</b> .....	61
5	<b>REFERENCIAS BIBLIOGRÁFICAS</b> .....	61

## LISTA DE TABELAS

---

<b>Capítulo 1</b>	
Tabela 1. Principais estudos nacionais com uso de RNAs para predição do valor genético .....	35
Tabela 2. Amostra de registros dos animais da raça Tabapuã contidos no banco de dados EMBRAPA GADO DE CORTE (MS).....	36
Tabela 3 – Valores da média e coeficiente de variação (CV%) de peso aos 205 dias (P205), peso ao nascimento (PNASC) e idade da mãe ao parto (IDMP).....	36
Tabela 4. Divisão do conjunto de dados inicial.....	37
Tabela 5. Propostas de utilização da RNA para predição do valor genético VG_P205_RNA .....	38
Tabela 6. Valores do erro quadrático médio de acordo com as propostas estabelecidas ....	39
Tabela 7A. Classificação baseada em valores genéticos de peso aos 205 dias preditos pelo BLUP (P205-BLUP) e RNA (P205-RNA) para animais da raça Tabapuã .....	41
Tabela 7B. Percentagens de valores classificatórios coincidentes e não coincidentes entre classificação obtida por meio do BLUP e RNA.....	41
Tabela 8. Correlação de Pearson (r) entre saídas BLUP e RNA, para animais inseridos no processo de validação após processo de treinamento ter sido finalizado .....	42
Tabela 9. Acuracias obtidas pelo algoritmo MOBJ .....	42
<b>Capítulo 2</b>	
Tabela 1. Diferentes composições genéticas dos animais em estudo.....	47
Tabela 2. Divisão do conjunto de dados inicial.....	49
Tabela 3. Propostas de utilização da RNA para predição da produção de leite em 305 dias .....	50
Tabela 4. Divisão do conjunto de dados inicial.....	52
Tabela 5. Propostas de utilização da RNA para predição da produção de leite em 305 dias .....	52
Tabela 6. Valores do erro quadrático médio (EQM) de acordo com as propostas estabelecidas.....	53
Tabela 7. Correlação entre a saída RNA e produção de leite em 305 dias de lactação de acordo com a participação de genes da raça Gir dos animais avaliados, para os dados de validação. ....	57
Tabela 8. Correlação entre a CPP e produção de leite em 305 dias de lactação de acordo com a participação de genes da raça Gir dos animais avaliados, para os dados de validação .....	57
Tabela 9. Correlação de Pearson entre a saída da RNA PL305_RNA e o valor real da PL305_2, de acordo com as propostas de montagem da RNA .....	57
Tabela 10. Correlação entre produção de leite em 305 dias e valor mensurado pela CPP e correlação entre a saída da RNA e valor obtido pelo CPP .....	58
Tabela 11. Estimativas do erro quadrático médio (EQM ou MSE) de acordo com diferentes propostas da RNA.....	58
Tabela 12. Exemplo de PL305_3 gerada pela RNA em comparação com PL305_3 original .....	60
Tabela 13. Correlação de Pearson entre a saída da RNA (PL305_RNA) e valor real da PL305_3, de acordo com as propostas de montagem da RNA .....	61

---

## LISTA DE FIGURAS

Figura 1. Componentes de um neurônio biológico (BRAGA, 2009).....	18
Figura 2. Neurônio de MCCULLOCH e PITTS (1943).....	19
Figura 3. Gráfico de função limiar .....	20
Figura 4. Gráfico função sigmóide.....	21
Figura 5. Gráfico função sigmóide com $a$ tendendo ao infinito.....	21
Figura 6. Gráfico função tangente hiperbólica .....	21
Figura 7. RNA de única camada (BRAGA et al, 2000) .....	22
Figura 8. RNA de múltiplas camadas (BRAGA et al, 2000) .....	22
Figura 9. Arquitetura de RNA recorrente (BRAGA et al, 2000) .....	22
Figura 10. Aprendizado supervisionado (BRAGA, 2007) .....	24
Figura 11. Fase de propagação .....	24
Figura 12: Etapa de retropropagação.....	25
Figura 13. Conjunto Pareto-ótimo.....	28
<b>Capítulo 1</b>	
Figura 1. Conexões neurais em uma RNA multicamada.....	37
Figura 2. Desempenho obtido pela RNA de acordo com o número de épocas .....	40
Figura 3. Predição dos valores genéticos para peso aos 205 dias por intermédio de RNA.....	40
<b>Capítulo 2</b>	
Figura 1. Esquema simplificado da RNA para predição da produção leite no segundo parto.....	49
Figura 2. Conexões neurais em uma RNA multicamada.....	49
Figura 3. Esquema simplificado da RNA para predição da produção leite no terceiro parto.....	51
Figura 4. Proposta A (EQM = 0.00152).....	53
Figura 5. Proposta B (EQM = 0.00011) .....	53
Figura 6. Proposta C (EQM = 0.00051) .....	54
Figura 7. Proposta D (EQM = 0.00001).....	54
Figura 8. Proposta E (EQM = 0.00001) .....	55
Figura 9. Proposta F para dados integrais sem ajuste ou restrições (EQM = 0.01085)...	55
Figura 10. Proposta D para dados completos ajustados (EQM = 0.00001).....	56
Figura 11. Proposta A após selecionar camada representativa de dados.....	59
Figura 12. Proposta B após selecionar camada representativa de dados.....	59
Figura 13. Proposta C após selecionar camada representativa de dados.....	60

## RESUMO

Foi aplicada a técnica de Redes Neurais Artificiais (RNA) para avaliação genética da característica peso corporal em animais Tabapuã e predição da produção total de leite em animais de diferentes composições genéticas. Em primeira análise, os resultados da RNA e BLUP apresentaram correlação positiva alta embora a classificação dos valores genéticos para peso aos 205 dias de idade (VG\_P205\_RNA) obtidos por meio da RNA e aqueles preditos pelo BLUP, demonstraram pequena variação, indicando riscos no uso das RNAs para avaliação genética da característica peso corporal em animais Tabapua. Posteriormente, inserções de novas informações no conjunto de dados sugeriram a necessidade de novo treinamento da RNA, sempre dependente do BLUP. Em última análise para predição da produção total de leite em animais de diferente composição genética, a capacidade provável de produção (CPP) foi calculada e comparada à produção total de leite observada em 305 dias de lactação e resultados obtidos pela RNA. Os resultados obtidos por intermédio da RNA demonstraram forte associação com a produção total observada e capacidade provável de produção (CPP), sendo a ultima forma simplificada para obtenção desta informação.

**Palavras-chave:** algoritmo Levenberg-Marquardt, algoritmo Backpropagation, redes neurais artificiais, avaliacao genetica

## ABSTRACT

The technique of Artificial Neural Networks (ANN) was used both for genetic evaluation of Tabapuã beef cattle body weight and for predicting total milk production of dairy cows of different genetic compositions. In the first analysis, the results from ANNs and BLUP were highly correlated although the ranked breeding values for body weight at 205 days (VG\_P205\_RNA) obtained by ANNs procedure and those predicted by BLUP, showed a small variation, indicating some risks in the use of ANNs for body weight genetic evaluation of Tabapuã cattle. Furthermore, insertions of new information in the data set suggested the ANNs procedure needs new training data, always dependent on BLUP. The last analysis to predict total milk production for dairy cows of different genetic composition, the milk producing ability (MPA) was calculated and compared with the observed milk production in 305 days of lactation and the results obtained by ANNs procedure. The results from ANNs procedure showed a strong association with the observed milk production and milk producing ability (MPA), the latter one being a simplified form to obtain this information.

**Keywords:** Levenberg-Marquardt algorithm, Backpropagation algorithm, artificial neural network, genetic evaluation

## 1. INTRODUÇÃO

O objetivo do melhoramento genético animal é modificar a base genética das diversas populações de animais. Deseja-se aumentar a frequência de alelos e de genótipos desejáveis e promover o incremento do mérito fenotípico médio de características importante economicamente. Para isto, a seleção e métodos de acasalamento são utilizados com base na busca de indivíduos que irão contribuir com alelos favoráveis para a próxima geração. A maior dificuldade em se identificar animais desejáveis é que o mérito genético dos animais representados pelo conjunto de genes (genótipos) não é informação obtida visivelmente, e sim, por meio de estimativas oriundas de avaliação genética (PEREIRA, 2004).

O valor genético aditivo dos membros da população é estimado a partir de informações fenotípicas, pré-ajustadas ou corrigido para fatores sistemáticos que interfiram no desempenho dos animais. As informações são oriundas do próprio indivíduo, de ancestrais e colaterais do indivíduo, suas progênes e combinações destas fontes.

O uso de resultados da avaliação genética em programas de seleção viabiliza e acelera os objetivos buscados em programas de melhoramento. Segundo CUNHA et al (2006), a pecuária bovina de corte brasileira demonstra elevados índices de desenvolvimento, responsável por parcela significativa da receita creditada ao agronegócio. Tal fato justifica pesquisas e implementação de programas de avaliação genética aplicados a diferentes raças, com o

objetivo de fornecer informações fidedignas ao crescimento do setor.

Na área de melhoramento animal, a metodologia de Modelos Mistos (Mixed Model - MM) proposta por HENDERSON (1949) tem sido utilizada como ferramenta padrão na avaliação de animais, em bovinocultura de leite, em bovinocultura de corte em outras espécies para obtenção do Melhor Preditor Linear Não Viesado (Best Linear Unbiased Predictor - BLUP). O ponto limitante da metodologia de HENDERSON é a alta demanda computacional relacionada à inserção de informações de todos os indivíduos presentes na base genealógica da população (NEVES, 2007).

O número total de equações a serem calculadas simultaneamente requer demasiado gasto computacional, elevados recursos de memória e processadores potentes, fatores que são exigidos ao extremo quando se trata de análises multicaracterística.

A abordagem de Redes Neurais Artificiais (RNAs) foi desenvolvida no final da década de 80 e início de 90, auxiliando estudos em diversas áreas de pesquisas.

Em melhoramento animal, as RNAs têm sido aplicadas de forma crescente em modelagem para predições de períodos de lactações em bovino leiteiro, demonstrando benefícios quando comparadas a outros métodos empregados (LACROIX et al.,1997) e, também, na predição de valores genéticos de bovinos de corte da raça Nelore (MEIRELLES, 2005).

As RNAs permitem a otimização de avaliações genéticas de características múltiplas e da inclusão de número maior de informações disponíveis (efeitos ambientais, manejo, rebanho, época de nascimento) por vez para classificar determinado animal. (NEVES, 2007).

ANDRADE (2009) analisou padrões cinemáticos da marcha em eqüinos por meio de RNAs. O algoritmo utilizado Multi-Objetivo LASSO alcançou resultados com alta generalização e topologias reduzidas. Os resultados mostram ainda que tal método foi capaz de eliminar entradas e realizar seleção automática de características de entrada importantes para a rede, sendo ferramenta importante para a escolha de características informativas para as RNAs.

Uma rede devidamente treinada não só deverá responder adequadamente aos padrões utilizados no processo de treinamento mas também aos demais que porventura sejam mostrados a ela. A esta propriedade de resposta coerente a padrões desconhecidos dá-se o nome de capacidade de generalização de uma rede (TEIXEIRA, 2001).

O uso de RNAs em avaliações genéticas merece atenção especial pelo fato da inserção de novos animais na avaliação necessitarem de novo treinamento da RNA. Por se tratar de um modelo classificatório de touros e o mercado de sêmen no cenário nacional absorve grande parte das vendas nos primeiros animais classificados, mínimas alterações na classificação afetará a contribuição genética de cada animal superior nos diversos rebanhos brasileiros.

Outros estudos também relataram a aplicabilidade e eficiência de RNAs em avaliações genéticas de bovinos de leite e corte (NEVES et al., 2007; MEIRELLES et. al., 2005; BERTAZZO, 2006).

Até o presente momento, não há registro de estudos na literatura brasileira que abordem a aplicação de RNAs em predições de lactações a partir de informações obtidas no dia do controle.

Considerando o fornecimento da alimentação baseado em dados de produções anteriores, nota-se que animais que apresentam maior declínio na produção de leite necessitam de maior suplementação de concentrado para manterem o mesmo nível de produção de vacas que apresentam menor declínio. Animais com moderada produção leiteira no período inicial da lactação, associada à maior persistência, apresentam menores esforços fisiológicos; elevando seu desempenho reprodutivo e proporcionando menores índices de doenças metabólicas, comparados aos das vacas de alta produção inicial e rápido declínio (MADSEN, 1975).

Observa-se no início da lactação um fator decisivo para a produção de leite. A maioria dos animais encontra-se em balanço energético negativo, caracterizado pela baixa ingestão de nutrientes por parte dos animais recém paridos, e ao mesmo tempo, por aumento contínuo da produção de leite. Dessa forma, vacas com elevadas produções de leite, em decorrência deste déficit de energia podem apresentar problemas como ocorrência de doenças metabólicas e fatores relacionados à reprodução.

As principais características estudadas referentes à permanência da fêmea no rebanho leiteiro, de acordo com DIAS (1997), são: longevidade (idade do animal ao descarte); duração da vida produtiva (intervalo do primeiro parto ao descarte); número de lactações durante a vida e habilidade de permanência no rebanho (HP) ou “stayability”.

Esta última pode ser medida em animais em produção, ao passo que as demais exigem a data de saída da fêmea do rebanho para a mensuração. ROBERTSON & RENDEL (1950) citaram algumas vantagens econômicas quando o animal permanece mais tempo no rebanho, tais como: (a) redução do custo anual das reposições; (b) aumento na produção média do rebanho; (c) redução do número de reposições (novilhas) a serem criadas, permitindo um aumento potencial do rebanho e (d) aumento da possibilidade de seleção de fêmeas. Em gado leiteiro, a preocupação dos produtores e pesquisadores está voltada aos efeitos do alto nível de produção das vacas sobre a vida útil no rebanho, sempre com intuito de minimizar os custos de produção (MARCONDES et al., 2005).

Segundo RIBEIRO (2001), o maior tempo de permanência do animal em produção tem sido associado à maior lucratividade em bovinos leiteiros. O investimento feito no animal, do nascimento até o primeiro parto, só é amortizado na primeira lactação, sendo necessário, pelo menos, duas lactações para que a vaca retorne o investimento realizado. Por isso, a saída do animal do rebanho está diretamente relacionada à lucratividade da fazenda leiteira, sendo que a produção de leite, a vida do

animal, a duração da vida produtiva e os custos relacionados à reposição do animal descartado devem ser fatores principais a serem observados quando da decisão do descarte (QUEIROZ & MCALLISTER, 2002).

Os critérios de descarte podem ser voluntários ou involuntários (VAN ARENDONK, 1996). O primeiro se deve à vontade do produtor e se relaciona com aspectos de produção, enquanto o descarte involuntário caracteriza-se por ser alheio à vontade do produtor e se deve, principalmente, a aspectos de sanidade. Segundo o autor, a diminuição das taxas de descartes involuntários levaria à maximização da longevidade e à otimização dos lucros, bem como permitiria ao produtor realizar maior taxa de descarte voluntário, aumentando o ganho genético.

Para SILVA et al. (2003), o fracasso reprodutivo é a principal causa de descarte das fêmeas em um rebanho, e a eficiência reprodutiva é o parâmetro de maior impacto econômico na atividade pecuária. Desta forma, a característica HP pode ser um dos critérios de seleção utilizados para aumentar o índice de fertilidade do rebanho (PARNELL, 2000).

GLÓRIA (2006) relatou a participação crescente do gado mestiço, obtido pelo cruzamento entre a raça Holandesa e raças Zebuínas, na produção de leite brasileira, por serem animais de genética mais rústica, com maior produção leiteira e melhor eficiência reprodutiva em ambientes tropicais e subtropicais. Entretanto, poucos estudos a respeito da predição das futuras lactações a partir de informações das lactações anteriores, em

diferentes composições genéticas, foram publicados. Tal informação pode ser de grande valia como auxílio nos critérios de descarte por parte dos produtores e técnicos da área de bovinocultura leiteira.

O objetivo desse estudo é avaliar o uso de RNAs na predição de lactações de animais Gir Leiteiro e cruzamentos com animais da raça Holandesa e em avaliações genéticas de touros da Raça Tabapuã para peso aos 205 dias.

Sendo assim, o presente estudo foi dividido em dois capítulos:

1. O primeiro capítulo trata da aplicação de Redes Neurais Artificiais na avaliação genética de animais da raça Tabapuã, com diferentes topologias da rede desenvolvida.
2. No segundo capítulo foi realizado estudo para predição da produção de leite no segundo parto por intermédio de informações do primeiro parto utilizando Redes Neurais Artificiais e predição da produção de leite no terceiro parto por meio de dados do primeiro e segundo partos.

## **2. REVISÃO DE LITERATURA**

O final da década de 80 marcou o ressurgimento da área de Redes Neurais Artificiais (RNAs), também denominadas como sistemas de processamento paralelo ou conexionismo. Tal forma de computação não algorítmica é caracterizada por sistemas que simulam a estrutura cerebral humana. Por não se basear em regras, a computação neural se constitui

como alternativa à metodologia de computação algorítmica convencional (BRAGA et al., 2007).

De acordo com o mesmo autor, RNAs são sistemas paralelos distribuídos compostos por unidades de processamento simples ou também denominados neurônios artificiais, que calculam determinadas funções matemáticas normalmente não lineares. Tais unidades são dispostas em uma ou mais camadas e interligadas por número elevado de conexões, em sua maioria unidirecional. Na maioria dos modelos essas conexões estão associadas a pesos, os quais armazenam o conhecimento adquirido pelo modelo e servem para ponderar a entrada recebida por neurônio da rede.

A solução de problemas por meio de RNAs é bastante atrativa, em razão da forma como estes são representados internamente pela rede e ao paralelismo natural inerente à arquitetura das RNAs que criam a possibilidade de desempenho superior ao dos modelos convencionais. Em RNAs, o procedimento usual na solução de problemas passa inicialmente por uma fase de aprendizagem em que o conjunto de exemplos ou base de dados é apresentado à rede, que extrai as características ou informações necessárias para representar a informação fornecida. Tais características são utilizadas posteriormente para gerar respostas para o problema.

A capacidade de aprendizado por meio de exemplos e de generalizar a informação aprendida é, sem dúvida, o atrativo principal da solução de

problemas por meio de RNAs. A generalização, que está associada à capacidade de a rede aprender por meio de um conjunto reduzido de exemplos e posteriormente dar respostas coerentes para dados não conhecidos, é uma demonstração de que a capacidade das RNAs vai muito além do que simplesmente mapear relações de entrada e saída. As RNAs são capazes de extrair informações não apresentadas de forma explícita por meio dos exemplos. Não obstante, as RNAs têm capacidade de atuar como mapeadores universais de funções multivariáveis, com custo computacional que cresce apenas linearmente com o número de variáveis. Outra característica fundamental é sua auto-organização e de processamento temporal que, vinculada àquelas citadas anteriormente, faz das RNAs ferramenta computacional atrativa para a solução de problemas complexos.

Dessa forma, as RNAs são indicadas para problemas cujas soluções requerem conhecimentos que sejam difíceis de especificar, mas para os quais existam dados ou observações suficientes (ZHANG et al., 1998). Esta capacidade de aprender com os próprios dados da série (modelagem empírica) pode apresentar problemas, pois nem sempre os relacionamentos entre as amostras são evidentes e as observações são freqüentemente mascaradas. Segundo (SHARKEY, 1999), para a previsão de séries temporais, os melhores resultados são obtidos pela combinação de diferentes modelos de previsão ao invés da seleção daquele que apresenta melhor desempenho individual.

RNAs nem sempre serão a melhor alternativa para descoberta e

reconhecimento de padrões. São fundamentais o conhecimento prévio do problema em questão e o relacionamento entre as variáveis abordadas.

Segundo BRAGA et. al (2007), o primeiro modelo artificial de um neurônio biológico foi fruto do trabalho pioneiro de Warren McCulloch e Walter Pitts, em 1943. Em trabalho publicado pelos autores no mesmo ano, denominado “A Logical Calculus of the Ideas Immament in Nervous Activity” são apresentadas discussões sofisticadas de redes lógicas de neurônios artificiais, denominados MCP (iniciais dos sobrenomes dos pesquisadores), além de novas idéias sobre máquinas de estados finitos, elementos de decisão de limiar lineares e representações lógicas de várias formas de comportamento de memória. Tal trabalho concentrou-se na descrição do modelo artificial de um neurônio e apresentar suas capacidades computacionais do que em apresentar técnicas de aprendizado. O aprendizado de redes biológicas e artificiais veio a ser objeto de estudo somente alguns anos depois.

HEBB (1949) apresentou o primeiro trabalho demonstrando ligação direta com o aprendizado de RNAs. Tal trabalho demonstrou como a plasticidade de aprendizagem das redes neurais é conseguida por meio da variação dos pesos de entrada dos neurônios, aprendizado dos neurônios biológicos baseados no reforço das ligações sinápticas entre neurônios excitados. Tal regra foi interpretada sob metodologia matemática como ponto de partida para vários algoritmos de aprendizado.

Os neurônios biológicos são divididos, de maneira simplificada e simbólica, em três seções: corpo celular, dentritos e axônio (Figura 1). Cada um dos três componentes possui função específica e ao mesmo tempo complementares. Os dentritos têm por função receber as informações, ou impulsos nervosos, oriundas de outros neurônios e conduzi-las até o corpo celular. Neste local, a informação é processada e novos impulsos são gerados.

Esses impulsos são transmitidos a outros neurônios, conduzidos por meio do axônio até os dentritos dos neurônios seguintes. O ponto de contato entre a terminação axônica de um neurônio e o dentrito de outro é denominado sinapse. É pelas sinapses que os neurônios se unem funcionalmente, formando as redes neurais biológicas. As sinapses

funcionam como válvulas e são capazes de controlar a transmissão de impulsos – isto é, o fluxo da informação – entre os neurônios da rede neural.

A figura 1 ilustra, de forma simplificada, os componentes do neurônio biológico. Os sinais oriundos dos neurônios pré-sinápticos são transmitidos para o corpo do neurônio pós-sináptico, no qual são comparados a outros sinais recebidos pelo mesmo. Caso o percentual de excitações do neurônio seja suficientemente alto, a célula produz impulso que é transmitido para as células seguintes (neurônios pós-sinápticos). Esse sistema simples é responsável pela maioria das funções realizadas pelo nosso cérebro. A capacidade de realizar funções complexas surge com a operação em paralelo de todos os neurônios do nosso sistema nervoso.

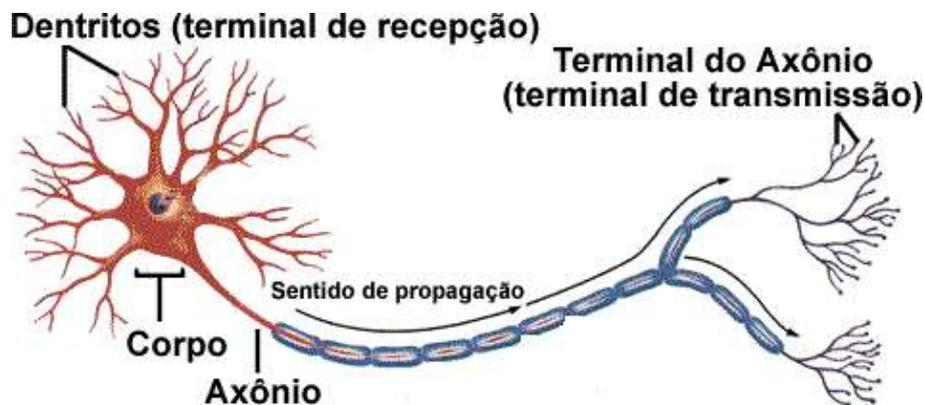


Figura 1. Componentes de um neurônio biológico (BRAGA, 2009).

O cérebro humano é responsável por funções cognitivas básicas, assim como pela execução de funções sensoriomotoras e autônomas. Além disso, sua rede de neurônios tem a

capacidade de reconhecer padrões e relacioná-los, usar e armazenar o conhecimento por experiência, além de interpretar observações. O comportamento individual de cada

neurônio biológico é bem entendido do ponto de vista funcional, sendo base para o desenvolvimento das RNAs. Estas reproduzem as funções básicas das redes biológicas, por implementações de seu comportamento funcional e de sua dinâmica. Como características comuns, pode-se citar que os dois sistemas são baseados em unidades de computação paralela e distribuída que se comunicam por meio de conexões sinápticas, possuem detectores de características, redundância e modularização das conexões.

O neurônio biológico pode ser visualizado do ponto de vista funcional da seguinte forma: as suas múltiplas entradas recebem ativações excitatórias ou inibitórias dos neurônios anteriores, e caso essa soma de excitações e inibições ultrapasse determinado limite, o neurônio emite impulsos nervosos. O neurônio artificial proposto por MCCULLOCH e PITTS (1943) nada mais é do que a simplificação do que se sabia na época à respeito do neurônio biológico. Sua descrição matemática resultou em modelo com  $n$  terminais de entrada (dentrito) que recebem os valores  $x_1, x_2, \dots, x_n$  (que representam as ativações dos neurônios anteriores) e apenas um terminal de saída  $y_j$  (representando o axônio). Para representar o comportamento das sinapses, os terminais de entrada do neurônio têm pesos acoplados  $w_1, w_2, \dots, w_n$ , cujos valores podem ser positivos ou negativos, dependendo de as sinapses correspondentes serem inibitórias ou excitatórias. O efeito de uma sinapse particular  $i$  no neurônio pós-sináptico é dado por  $x_i w_i$ . Os pesos determinam o grau no qual o neurônio deverá considerar sinais de disparo que ocorrem

naquela conexão. A figura 2 ilustra o modelo descrito acima, sendo  $\Sigma$  a soma ponderada das entradas e  $f(\cdot)$  a função de ativação.

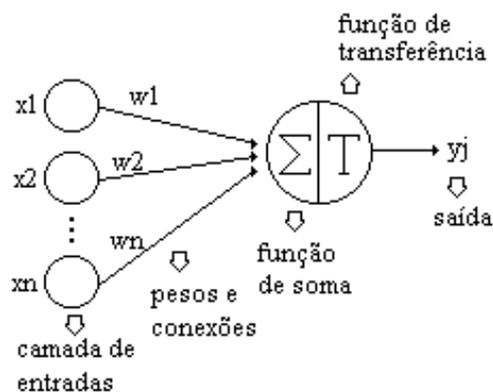


Figura 2. Neurônio de MCCULLOCH e PITTS (1943).

No modelo biológico, o neurônio dispara quando a soma dos impulsos que ele recebe ultrapassa o seu limiar de excitação (*threshold*). Esse comportamento do neurônio biológico, por sua vez, é representado no modelo artificial por um mecanismo simples, que faz a soma dos valores  $x_i w_i$  recebidos pelo neurônio (soma ponderada) e decide se o neurônio deve ou não disparar (saída igual a 1 ou 0), comparando a soma obtida ao limiar ou *threshold* do neurônio. No modelo MCP, a ativação do neurônio é obtida por meio da aplicação da função de ativação, que ativa ou não a saída, dependendo do valor da soma ponderada das suas entradas. (BRAGA et. AL, 2000)

A saída do neurônio é definida pelo seu valor de ativação calculado da seguinte forma:

$$v_j = \sum_{i=1}^m w_{ji} x_i + b$$

em que:

$v$  é o valor de ativação do neurônio  $k$ ;

$w$  são os pesos das conexões do neurônio  $k$ ;

$x$  é o valor de cada um dos  $m$  estímulos que chegam ao neurônio  $k$ ;

$b$  é o valor do bias que será somado ao valor do combinador linear para compor o valor de ativação.

### Funções de ativação

A função de ativação é responsável pela geração da saída  $y_j$  do neurônio a partir dos valores dos vetores de peso  $w = \{w_1, w_2, \dots, w_n\}$  e de entrada  $x = \{x_1, x_2, \dots, x_n\}$ . A função de ativação de um neurônio MCP anterior é do tipo degrau deslocada do limiar de ativação  $\Theta$  em relação à origem, ou seja, a saída será 1 ou 0.

A seguir, são apresentados três tipos básicos de funções de ativação comumente utilizadas nos modelos RNAs.

#### 2.1 Função limiar (Degrau):

$$f(v) = \begin{cases} 1, & \text{se } v \geq 0; \\ 0, & \text{se } v < 0; \end{cases}$$

Neurônios construídos com essa função apresentam saída  $y = 0$  caso o valor de ativação  $v$  seja negativo e 1 nos casos em que o valor de ativação seja positivo (Figura 3).

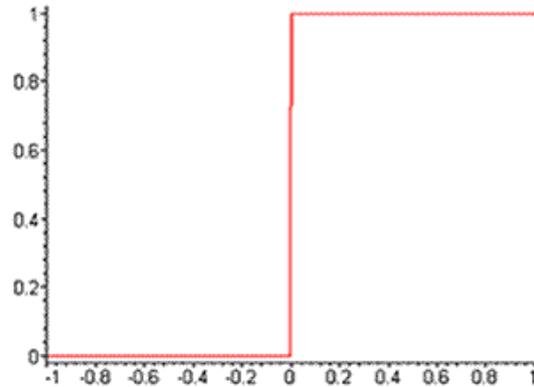


Figura 3. Gráfico de função limiar.

O valor de ativação ( $v$ ) é composto pelo combinador linear e pelo bias:

$$v_j = \sum_{i=1}^m w_{ji} x_i + b$$

#### 2.2 Função Sigmóide:

Esta função, ilustrada por meio da figura 4, ao contrário da função limiar, pode assumir todos os valores entre 0 e 1. É a forma mais empregada na construção de RNAs e sua representação mais utilizada é a função logística, definida por:

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

em que  $a$  é o parâmetro de inclinação da função sigmóide e  $v$  é o valor de ativação do neurônio.

Quando se aumenta o valor do parâmetro  $a$ , tendendo-o ao infinito, esta função comporta-se como uma função de limiar, como é ilustrado na figura 5. Observa-se que a função sigmóide é diferenciável, enquanto a função de limiar não possui esta propriedade.

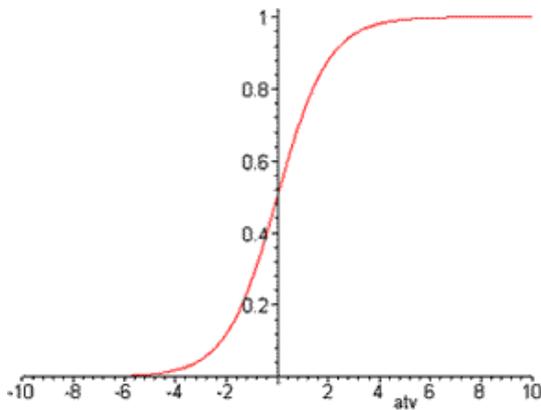


Figura 4. Gráfico função sigmóide.

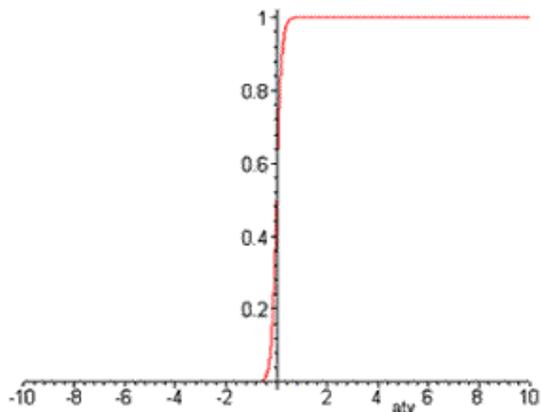


Figura 5. Gráfico função sigmóide com  $a$  tendendo ao infinito.

O fato da função de ativação tipo sigmóide assumir valores negativos permite que o aprendizado do neurônio se efetue de forma mais rápida, reduzindo o número de iterações de treinamento.

### 2.3 Tangente Hiperbólica:

Como a Função logística, também possui forma de "s", assumindo valores entre 1 e -1, sendo representada por:

$$f(v) = a \frac{e^{(bv)} - e^{(-bv)}}{e^{(bv)} + e^{(-bv)}},$$

em que:

$a$  é o parâmetro de inclinação da curva;  
 $b$  são os limites inferiores e superiores ( $b = |1|$  no gráfico);  
 $v$  é o valor de ativação da unidade.

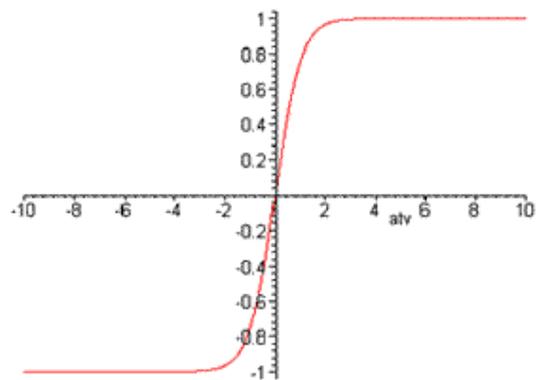


Figura 6. Gráfico função tangente hiperbólica.

Os neurônios individuais possuem capacidade computacional limitada independente da função de ativação utilizada no processo. No entanto, o conjunto de neurônios artificiais conectados na forma de RNAs é capaz de resolver problemas de complexidade elevada. (BRAGA et. al ,2007).

As figuras 7 e 8 ilustram possíveis configurações de neurônios artificiais conectados na forma de RNAs. A estrutura mais simples é apresentada na Figura 7, que corresponde a rede neural de camada única alimentada para frente (*feedforward*). Tais estruturas são capazes de resolver problemas multivariáveis de múltiplas funções acopladas, mas com algumas restrições

de complexidade, por serem de única camada. A figura 8 assemelha-se à figura 7, entretanto, possui uma camada adicional. A camada intermediária confere à RNA maior capacidade computacional e universalidade na aproximação de funções contínuas. As figuras 7 e 8 são consideradas estáticas, já que não apresentam recorrência em sua estrutura, sendo suas saídas dependentes apenas das entradas atuais. Caso isso não ocorra, existindo recorrência de conexões entre neurônios de um mesmo nível ou entre neurônios de saída e de camadas anteriores, denomina-se tal arquitetura de RNAs recorrentes (Figura 9).

A definição da estrutura de uma RNA para a resolução de determinado problema depende de vários fatores tais como: complexidade do problema, dimensionamento do espaço de entrada, características dinâmicas ou estáticas, conhecimento a priori sobre o problema e representatividade dos dados.

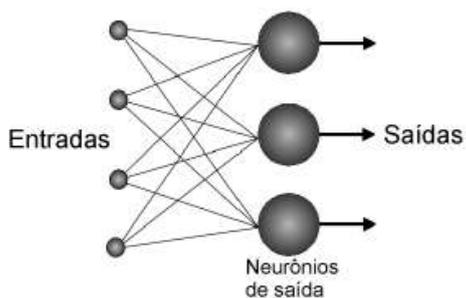


Figura 7. RNA de única camada (BRAGA et al, 2000).

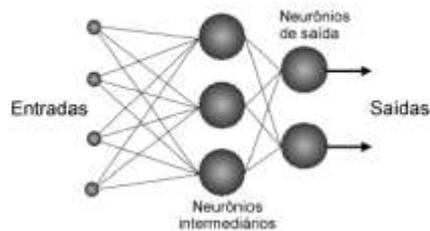


Figura 8. RNA de múltiplas camadas (BRAGA et al, 2000).

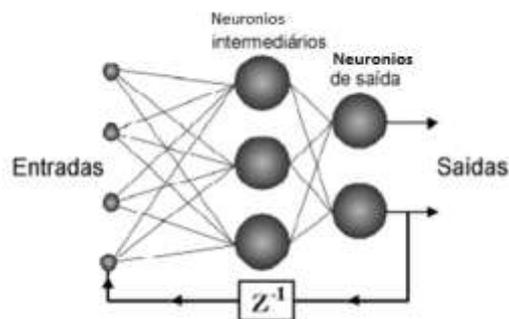


Figura 9. Arquitetura de RNA recorrente (BRAGA et al, 2000).

A propriedade mais importante das redes neurais é a habilidade de reconhecer padrões em seu ambiente e com isso melhorar seu desempenho. Isso é feito por meio de processo iterativo de ajustes aplicado a seus pesos, o treinamento. O aprendizado ocorre quando a rede neural atinge solução generalizada para uma classe de problemas.

Na abordagem conexionista o conhecimento não é adquirido por meio de regras explícitas como ocorrido em processos de Inteligência Artificial (IA) simbólicos, mas sim por meio de ajustes das intensidades de conexões entre os neurônios.

Segundo (HAYKIN, 2001), aprendizado é o processo pelo qual os parâmetros

livres de uma rede neural são ajustados por meio de uma forma continuada de estímulo pelo ambiente externo, sendo o tipo específico de aprendizado definido pela maneira particular como ocorrem os ajustes dos parâmetros livres. É importante ressaltar que o conceito de aprendizado está relacionado à melhoria do desempenho da rede segundo algum critério preestabelecido. O erro quadrático médio da resposta da rede em relação ao conjunto de dados fornecido pelo ambiente é utilizado como critério de desempenho pelos algoritmos de correção de erros. Quando estes algoritmos são utilizados em treinamentos da RNA, estima-se que o erro diminua à medida que o aprendizado prossiga.

Segundo BRAGA (2007), há vários algoritmos diferentes para treinamento das RNAs, os mesmos podem ser agrupados em dois paradigmas: aprendizado supervisionado e não-supervisionado. Estes algoritmos diferem entre si principalmente pelo modo como os pesos são modificados.

Outro fator importante é a maneira pela qual uma rede neural se relaciona com o ambiente. Nesse contexto existem os seguintes paradigmas de aprendizado:

- **Aprendizado supervisionado**, quando é utilizado um agente externo que indica à rede a resposta desejada para o padrão de entrada;
- **Aprendizado não supervisionado** (auto-organização), quando não existe um agente externo indicando a resposta desejada para os padrões de entrada;

No aprendizado supervisionado pares de entrada e saída, chamados de conjunto de treinamento, são apresentados à rede

não treinada. O erro calculado, que representa a diferença entre o valor gerado e o desejado, é utilizado para ajustar os pesos entre os neurônios com o objetivo de aproximar a resposta da rede à saída alvo (JAIN et al, 1996).

Essa iteratividade do treino deverá ser repetida com todo o conjunto de treinamento, de modo a minimizar o erro da rede até um valor aceitável quando a rede é dita treinada. Após a fase de treinamento, é realizada a fase de teste com dados (conjunto de teste) que ainda não foram apresentados à rede. Dessa forma, assegura-se que a rede apenas não memorizou o conjunto de treinamento (*overfitting*), mas aprendeu os padrões gerais envolvidos na aplicação (generalização).

A figura 10 ilustra a representação esquemática do aprendizado supervisionado. Para cada padrão de entrada, a rede tem sua saída corrente comparada à saída desejada pelo supervisor, que fornece informações sobre a direção de ajuste dos pesos. A minimização da diferença é incremental, já que pequenos ajustes são feitos nos pesos a cada etapa de treinamento, de tal forma que estes caminhem (caso haja solução possível) para uma solução. Os exemplos mais conhecidos de aprendizado supervisionado são a regra Delta (WIDROW & HOFF, 1960) e sua generalização para redes de múltiplas camadas, o algoritmo Backpropagation (RUMELHART et al; 1986).

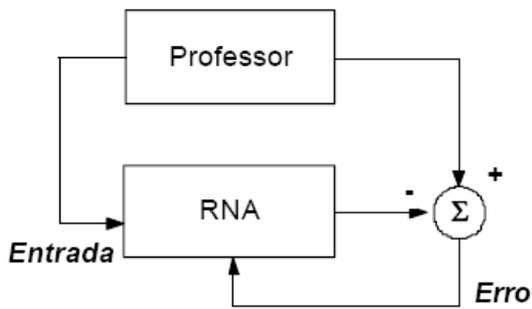


Figura 10. Aprendizado supervisionado (BRAGA, 2007).

Existem duas formas de implantação do aprendizado supervisionado, denominadas off-line e on-line. Na primeira, os dados do conjunto de treinamento não variam, e, uma vez obtida solução para a rede, esta deve permanecer fixa. Caso novos dados sejam adicionados, um novo treinamento, envolvendo também os dados anteriores, deve ser realizado para evitar interferência do treinamento anterior. Em aprendizados on-line, o conjunto de dados muda continuamente, e a rede deve estar em processo contínuo de adaptação.

No aprendizado não supervisionado, apenas as entradas são apresentadas à rede e esta explora a estrutura interna nos dados, ou correlações entre os modelos nos dados, e organiza padrões dentro de categorias a partir dessas correlações (JAIN et al, 1996).

Segundo BRAGA et al (2000), backpropagation é o algoritmo para treinamento de redes multicamadas mais difundido e baseia-se no método de aprendizado supervisionado por correção de erros. A metodologia do algoritmo é abaixo descrita:

**1º - Etapa de propagação:** Conforme ilustração da figura 11, após apresentação do padrão de entrada à RNA, a resposta de uma unidade é propagada como entrada para as unidades na camada seguinte, alcançando a camada de saída, onde é obtida a resposta da rede e o erro é assim calculado;

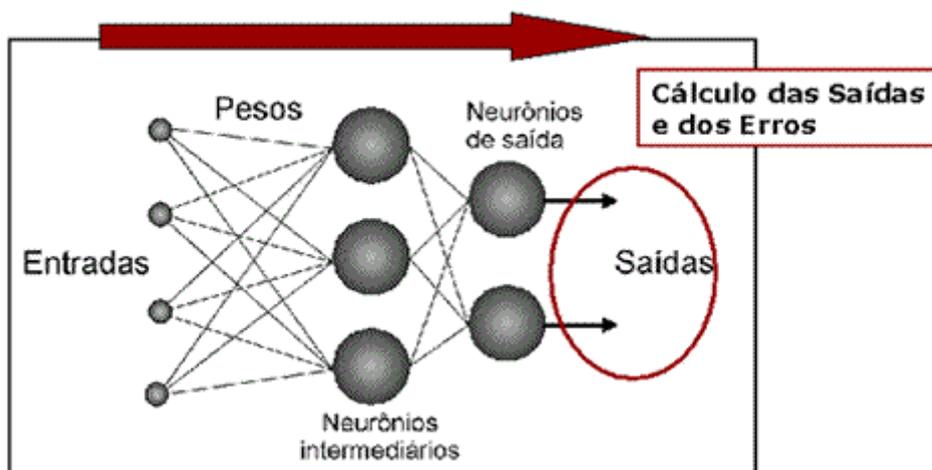


Figura 11. Fase de propagação.

2° - **Etapa de retropropagação ("backpropagation"):** Nesta etapa, os pesos sinápticos são alterados no sentido camada de saída em direção à camada de entrada. (Figura 12).

Durante a fase treinamento devem-se apresentar conjuntos formados pelo par:

entrada para a rede e valor desejado para resposta a entrada. A saída será comparada ao valor desejado (*target*) e será computado o erro global da rede, que influenciará na correção dos pesos no passo de retropropagação.

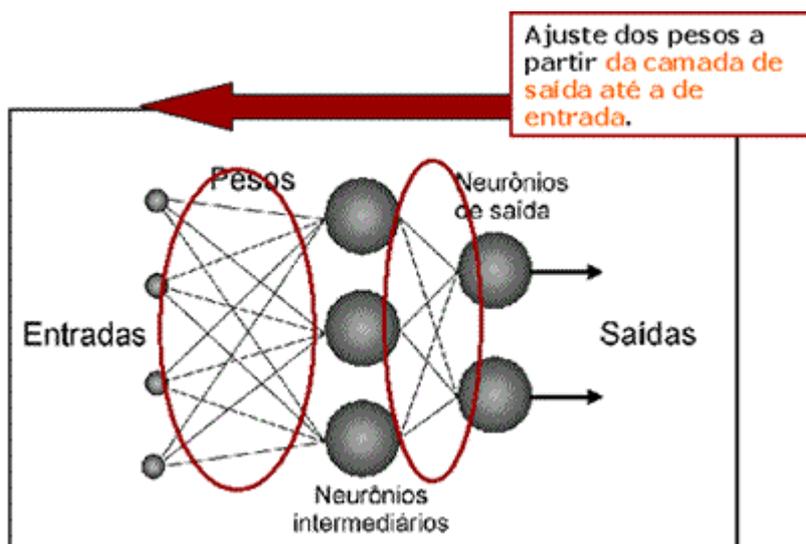


Figura 12: Etapa de retropropagação.

#### 2.4 Resumo do Algoritmo Backpropagation:

A fase forward envolve os seguintes passos:

1. O vetor de entrada  $\mathbf{x}$  é apresentado às entradas da rede, e as saídas dos neurônios da primeira camada escondida  $C_1$  são calculadas.
2. As saídas da camada escondida  $C_1$  proverão as entradas da camada seguinte  $C_2$ . As saídas da camada  $C_2$  são então calculadas. O processo se repete até que se chegue à saída  $C_k$ .

3. As saídas produzidas pelos neurônios da camada de saída são então comparadas às saídas desejadas  $y_d$  para aquele vetor de entrada  $\mathbf{x}$ , e o erro correspondente  $y_d - y$  é calculado.

O objetivo da fase forward é obter o erro de saída após a propagação do sinal por todas as camadas da rede. A fase backward é assim detalhada:

1. O erro da camada de saída  $C_k$  é utilizado para ajustar diretamente os seus pesos, fazendo uso do gradiente descendente do erro.

2. Os erros dos neurônios da camada de saída  $C_k$  são propagados para a camada anterior  $C_{k-1}$ , utilizando-se para isso os pesos das conexões entre as camadas, que serão multiplicados pelos erros correspondentes. Assim, tem-se um valor de erro estimado para cada neurônio da camada escondida que representa uma medida da influencia de cada neurônio da camada  $C_{k-1}$  no erro de saída da camada  $C_k$ .
3. Os erros calculados para os neurônios da camada  $C_{k-1}$  são então utilizados para ajustar os seus pesos pelo gradiente descendente, analogamente ao procedimento utilizado para a camada  $C_k$ .
4. O processo se repete até que os pesos da camada  $C_1$  sejam ajustados, concluindo o ajuste dos pesos de toda a rede para o vetor de entrada  $x$  e sua saída desejada  $y_d$ .

Segundo WIDROW & HOFF (1960), o ajuste dos pesos feito pelo algoritmo backpropagation baseia-se na regra delta proposta pelos mesmos autores. A generalização desse algoritmo para redes multi-camadas é conhecido por delta generalizada ou backpropagation. Nesse caso, a função de custo a ser minimizada é a função de erro definida pela soma dos erros quadráticos.

Existem diversos algoritmos para treinamento de RNAs, sendo alguns implementados com objetivo de minimizar os erros quadráticos (*como exemplo Backpropagation, Quickprop, Levenberg Marquardt e Rprop*) e outros com funcionalidade adicional de maximizar a capacidade de generalização como algoritmos de *Pruning* (REED, 1993). A técnica de

*pruning* ou poda envolve a eliminação de pesos e neurônios irrelevantes para a função executada pela rede. Dois métodos são conhecidos, um baseado na avaliação de sensibilidade da saída e o outro envolvendo modificações na função de custo. No primeiro grupo, a verificação da variação do erro é feita após retirada dos elementos da rede. Caso a retirada do elemento não cause variação excessiva do erro de saída, conclui-se que a rede apresenta pouca sensibilidade ao elemento e este poderá ser retirado sem afetar a capacidade de modelagem dos dados. Entretanto, a correlação entre tais elementos não é avaliada no processo de *pruning*.

Segundo ANDRADE (2009), algoritmos de treinamento de redes neurais que utilizam apenas o conjunto de treinamento e o erro relacionado a ele para ajuste dos pesos não são sempre capazes de encontrar solução com alta capacidade de generalização. Estes algoritmos são muito sensíveis ao número de parâmetros livres,  $n_w$ , aos valores de inicialização dos mesmos e também ao instante de parada do treinamento, como é o caso dos algoritmos *Backpropagation, Quickprop, Rprop, Levenberg-Marquardt* e treinamento por modos deslizantes. De acordo com os parâmetros, as redes treinadas podem ficar sub ou super ajustadas aos dados de treinamento, prejudicando sua capacidade de generalização.

## 2.5 Algoritmo Levenberg-Marquardt

Enquanto o *back-propagation* padrão utiliza a descida de gradiente como método de aproximação do mínimo da

função erro, o algoritmo de Levenberg-Marquardt (LM) utiliza uma aproximação pelo método de Newton [2]. Esta aproximação é obtida a partir da modificação do método de Gauss-Newton introduzindo-se o parâmetro  $m$ , conforme a equação:

$$\Delta x = [J^T(x)J(x) + \mu I]^{-1} J^T(x)e(x)$$

em que  $I$  é a matriz identidade,  $e(x)$  é o erro e  $J$  é a matriz Jacobiana. O parâmetro  $m$  é multiplicado por um fator ( $b$ ) toda vez que um passo resultar num aumento na função erro, que se deseja minimizar. Quando um passo resultar na diminuição da função genérica  $V(x)$ ,  $m$  é dividido pelo fator  $b$ . Isto significa dizer que, se houver uma convergência para o mínimo da função,  $m$  é pequeno e o algoritmo se aproxima do método de Newton (passo  $1/m$ ); no caso de não haver convergência, o método se aproxima da descida de gradiente. O parâmetro  $m$  funciona como um fator de estabilização do treinamento, ajustando a aproximação de forma a utilizar a rápida convergência do método de Newton e evitando passos muito grandes que possam levar a um erro de convergência.

O algoritmo é uma técnica de otimização numérica que demanda grande quantidade de memória e elevada complexidade computacional, o que pode tornar sua utilização ineficiente para redes de tamanho elevado. (SOARES & NADAL, 1999)

## 2.6 Algoritmo Multi-Objetivo

Soluções que se apresentam super-ajustadas aos dados são soluções com

complexidade além da necessária e, portanto, possuem normas mais elevadas que a necessária. Por outro lado, soluções de normas baixas são aquelas que se apresentam sub-ajustadas ao conjunto de treinamento, ou seja, possuem baixa complexidade. Como a norma do vetor de pesos é uma medida de complexidade da função mapeada pela rede, surge a idéia da utilização desta medida como função custo a ser minimizada em conjunto com o erro em relação aos padrões de treinamento. Portanto, pode-se formalizar um problema de otimização multi-objetivo onde são considerados dois objetivos a serem minimizados.

Por meio da otimização multi-objetivo e da utilização das funções norma do vetor de pesos e erro de treinamento como funções de custo, pode-se gerar soluções de complexidades diferentes entre os extremos de baixa e alta complexidade, de forma que uma destas soluções tenha a complexidade adequada e, portanto, com as características desejáveis. Ao processo de geração de soluções e a escolha da solução final dá-se o nome de método MOBJ (Multi-Objetivo) (ANDRADE, 2009).

A proposta original do Algoritmo MOBJ consiste em otimizar as funções de erro e da norma (TEIXEIRA et AL., 2001). O conjunto Pareto representa o conjunto de soluções entre dois extremos: as soluções de mínimo erro (e alta norma) e a solução com mínima norma (e alto erro).

As duas etapas do treinamento MOBJ consistem, primeiramente, na geração do conjunto de soluções Pareto-ótimo e posteriormente na obtenção da melhor

solução de rede, presente nesse conjunto, através de uma regra de decisão.

A figura 13 apresenta um esquema da curva do Pareto-ótimo. As soluções que

pertencem à curva não podem ser melhoradas considerando os dois objetivos simultaneamente. O conjunto viável é aquele que se encontra acima da curva apresentada.

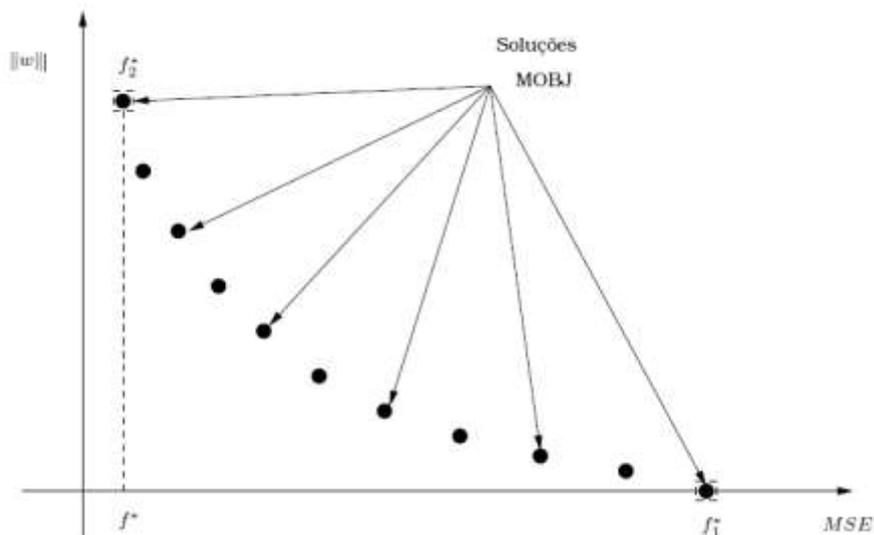


Figura 13. Conjunto Pareto-ótimo.

A figura 13 também ilustra soluções  $f_1^*$  e  $f_2^*$  nos extremos da curva. A solução  $f_1^*$  corresponde à solução sub-ajustada, que pode ser obtida igualando os pesos a zero. Já a solução  $f_2^*$  corresponde à solução sobre-ajustada, que pode ser obtida pelo treinamento da rede por um algoritmo padrão como o *Backpropagation*. As soluções intermediárias são geradas pelo algoritmo MOBJ. A solução  $f^*$  é a solução ideal (norma zero e erro mínimo) e que nunca pode ser obtida.

O algoritmo MOBJ resolve o problema de otimização multi-objetivo por meio da transformação desse problema em outro, mono-objetivo, onde os múltiplos objetivos aparecem como restrições. O segundo passo do método MOBJ é a escolha da melhor solução, dentre

aquelas presentes no conjunto Pareto-ótimo, conforme a equação:

$$\vec{w}^* = \arg \min e_V$$

em que:

$$e_V = \frac{1}{N_V} \sum_{i=1}^{N_V} [y_{V_i} - f(\vec{x}_{V_i}; \vec{w})]^2$$

$$\vec{w} \in \mathcal{W}^*$$

$$\mathcal{X} = \{\vec{x}_i, y_i\}_{i=1}^{N_T}$$

$N$  é o tamanho da amostra

$\mathcal{W}^*$  é o conjunto de soluções Pareto-ótimas para um dado conjunto de treinamento

### 3.REFERENCIAS BIBLIOGRAFICAS

ANDRADE, A. G. P. *Análise do padrão cinemático da marcha em eqüinos por meio de redes neurais artificiais*. 2009. 85f. Tese (Doutorado) - Escola de Veterinária, Universidade Federal de Minas Gerais, Belo Horizonte, MG.

BERTAZZO, R. P. *Utilização de redes neurais artificiais na predição do mérito genético de bovinos de corte*. Tese (Doutorado) - Faculdade de Zootecnia e Engenharia de Alimentos, Universidade Federal de Minas Gerais, Pirassununga, SP, 2006.

BRAGA, A. P.; CARVALHO, A. P. L. F.; LUDERMIR, T. B. *Redes neurais artificiais: teoria e aplicações*. Rio de Janeiro: LTC, 2000.

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

CUNHA, E.E.; EUCLYDES, R.F.; TORRES, R.A. *et al.* Simulação de dados para avaliação genética de rebanhos de gado de corte. *Arquivo Brasileiro de Medicina Veterinária Zootecnia*, v.58, p.381-387, 2006.

DIAS, A. S. C. *Estudo das características produtivas e de permanência em um rebanho de bovinos da raça Caracu*. 1997. 65f. Dissertação (Mestrado em Zootecnia) – Faculdade de Ciências Agrárias e Veterinárias, Universidade Estadual Paulista, Jaboticabal, SP.

GLÓRIA, J.R.; BERGMANN, J.A.G.; REIS, R.B.; COELHO, M.S.; SILVA,

M.A. Efeito da composição genética e de fatores de meio sobre a produção de leite, a duração da lactação e a produção de leite por dia de intervalo de partos de vacas mestiças Holandês-Gir. *Arquivo Brasileiro de Medicina Veterinária e Zootecnia*, v.58, n.6, p.1139-1148, 2006.

HAYKIN, S. *Redes neurais: princípios e prática*. Porto Alegre: RS: Bookman, 2001.

HEBB, D. O. *The organization of behavior: a neuropsychological theory*. New York: John. Wiley and Sons, 1949. 335p.

HENDERSON, C. R. S. Estimation of changes in herd environment. *Journal of Dairy Science*, v.32, n.8, p.709, 1949. (Abstract).

JAIN, A. K.; MAO, J.; MOHIUDDIN, K. M. Artificial neural networks: a tutorial. *Computer*, v.9, n.3, p. 31-44, 1996.

LACROIX, R.; SALEHI, F.; YANG, X. Z.; WADE, K. M. Effects of data preprocessing on the performance of artificial neural networks for dairy yield prediction and cow culling classification. *American Society of Agricultural Engineers*, v.40, n.3, p.839-846, 1997.

MADSEN, O. A comparison of some suggested measures of persistency of milk yield in dairy cows. *Animal Production*, v.20, p.191-197, 1975.

MARCONDES, C. R.; PANETO, J. C. C.; SILVA, J. A. *et al.* Comparação entre análise para permanência no rebanho de vacas Nelore utilizando modelo linear e modelo de limiar.

*Arquivo Brasileiro de Medicina Veterinária Zootecnia*, v.57, n.2, p.234-240, 2005.

MCCULLOCH, W.; PITTS, W. A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics*, v.7, p.115 – 133, 1943.

MEIRELLES, F. D. P. *Modelo computacional de um rebanho bovino de corte virtual utilizando simulação Monte Carlo e redes neurais artificiais*. 2005. 104f. Tese (Doutorado em Zootecnia) – Faculdade de Zootecnia e Engenharia de Alimentos, Universidade de São Paulo, Pirassununga, SP.

NEVES, A. R. M. *Aplicação de redes neurais artificiais na predição de valores genéticos em bovinos de leite da raça pardo-suíça*. Belém: UFP, 2007. 74p.

PARNELL, P. Balancing growth, carcass and fertility in your breeding program. In: THE ANNUAL FEEDER STEER SCHOOL, 5., 2000. Armidale. *Proceedings...* Armidale: Feeder Steer School, 2000. p.90-96.

PEREIRA, J. C. C. *Melhoramento genético aplicado à produção animal*. Belo Horizonte: FEP-MVZ, 2004. 609p.

QUEIROZ, S. A.; MCALLISTER, A. J. Avaliação do tipo de descarte de vacas sobre a rentabilidade de rebanhos leiterios no estado de Kentucky. In: REUNIÃO ANUAL DA SOCIEDADE BRASILEIRA DE ZOOTECNIA, 33., 1996, Fortaleza, CE. *Proceedings...* Fortaleza: SBZ, 1996.

REED, R. Pruning algorithms: a survey.

*IEEE transactions on Neural Networks*, v.4, n.5, p.740-746, 1993.

RIBEIRO, A. C. *Avaliação genético-econômica de rebanho da raça Holandesa*. 2001. 125f. Tese (Doutorado em Produção Animal) – Faculdade de Ciências Agrárias e Veterinárias – Universidade Estadual Paulista, Jaboticabal, SP.

ROBERTSON, A.; RENDEL, J. M. The use of progeny testing with artificial insemination in dairy cattle. *Journal of Genetics*, v.50, n.1, p.21-31, 1950.

RUMELHART, D. E.; HINTON, G. E.; WILLIAMS, R. J. Learning internal representations by error propagation. In: RUMELHART, D.; MCCLELLAND, J. (Eds.). *Parallel distributed processing*. Cambridge: M.I.T. Press, 1986. v. 1, p. 318-362.

SHARKEY, A. J. C. *Combining artificial neural nets: ensemble and modular multi-net systems*. London, UK: Springer Verlag, 1999.

SILVA, J. A.; ELER, J. P.; FERRAZ, J. B. S.; OLIVEIRA, H. N. Análise genética da habilidade de permanência em fêmeas da raça Nelore. *Revista Brasileira de Zootecnia*, v.32, n.3, p.598-604, 2003.

SOARES, P. P. S.; NADAL, J. Aplicação de uma rede neural feedforward com algoritmo de Levenberg-Marquardt para classificação do segmento ST do eletrocardiograma. In: CONGRESSO BRASILEIRO DE REDES NEURAIAS, 4., 1999, São José dos Campos. *Anais...* São José dos Campos: ITA, 1999. p.384-389.

TEIXEIRA, R. A.; BRAGA, A. P.; TAKAHASHI, R. H. C.; SALDANHA, R. R.. Recent advances in the MOBJ algorithm for training artificial neural networks. *International Journal of Neural Systems*, v.11, n.3, p.265-269, 2001.

VAN ARENDONK, J. A. M. Economic importance and possibilities for improvement of dairy cows herd life. In: WORLD CONGRESS OF GENETIC APPLIED TO LIVESTOCK PRODUCTION, 3., 1986, Lincoln. *Proceedings...*Lincoln: University of

Nebraska Board of Regents, 1986. p. 95-100.

WIDROW, B.; HOFF, M. E. Adaptive switching circuits. *IRE WESCON Convention Record*, pt.4, p.96-104, 1960.

ZHANG, G. P., PATUWO, B. E.; HU, M. Y. (1998). Forecasting with artificial neural networks: The state of the art. *International Journal of Forecasting*, v.14, n.1, p. 35-62, 1998.

## CAPITULO 1

### USO DE REDES NEURAIAS ARTIFICIAIS NA PREDIÇÃO DE VALORES GENÉTICOS PARA PESO AOS 205 DIAS EM BOVINOS DA RAÇA TABAPUÃ

#### RESUMO

Dados de 19240 animais Tabapuã, provenientes de 152 fazendas localizadas em diversos estados brasileiros, nascidos entre 1976 e 1995, foram utilizados para predição do valor genético do peso aos 205 dias de idade (VG\_P205) por meio de Redes Neurais Artificiais (RNAs) e usando o algoritmo LM - Levenberg Marquardt para treinamento dos dados de entrada. Por se tratar de rede com aprendizado supervisionado, foram utilizados como saída desejada, os valores genéticos preditos pelo BLUP para a característica P205. Os valores genéticos do P205 obtidos pela RNA e os preditos pelo BLUP foram altamente correlacionados. A ordenação dos valores genéticos do P205 oriundos das RNAs e os valores preditos pelo BLUP (VG\_P205\_RNA) sugeriram que houve variação na classificação dos animais indicando riscos no uso de RNAs para avaliação genética desta característica. Inserções de novos animais necessitam de novo treinamento dos dados, sempre dependentes do BLUP.

**Palavras-chave:** Redes neurais artificiais, avaliação genética, melhor preditor linear não viesado

#### ABSTRACT

Data from 1940 Tabapuã animals from 152 farms located in different states of Brazil, born from 1976 to 1995, were used to predict genetic value of body weight at 205 days of age (BV\_P205) of Tabapuã beef cattle using Artificial Neural Networks (ANN) and LM algorithm - Levenberg Marquardt training for data entry. Due to the use of networks with supervised learning the predicted breeding values for P205 from BLUP were used as desired output. The breeding values for P205 obtained from RNA and those predicted by BLUP were highly correlated. The ranked breeding values for body weight at 205 days from RNA and those predicted by BLUP (VG\_P205\_RNA) showed a variation in the classification of animals indicating risks in the use of ANNs procedure for genetic evaluation of this trait. Insertions of new animals need new training data, always dependent on BLUP.

**Keywords:** Artificial neural networks, genetic evaluation, best linear unbiased predictor.

## 1 - INTRODUÇÃO

Segundo PEREIRA (2004), o objetivo do melhoramento genético animal é alterar geneticamente populações de animais. Faz-se isto no intuito de aumentar a frequência de alelos e de genótipos desejáveis, com o consequente reflexo favorável no mérito fenotípico médio de características destas populações que sejam importantes economicamente. Diante desse fato, o melhorista possui duas estratégias para alterar populações geneticamente, a seleção e o método de acasalamento.

No Brasil, a partir do final dos anos 80 começaram a surgir vários grupos independentes de avaliação genética para bovinos de corte e de leite. Estes grupos têm publicado sumários de touros com lista do valor genético aditivo para várias características. São calculados os valores genéticos aditivos dos animais dos diferentes rebanhos. O uso desses resultados nos programas de seleção viabiliza e acelera os objetivos preestabelecidos do programa de melhoramento.

Segundo KEPLER FILHO (1999), as metodologias para predição de valores genéticos fazem uso de fontes de informação indireta sobre o genótipo de cada animal. Como essas fontes são sempre baseadas em valores fenotípicos, não é possível conhecer com precisão os valores genéticos. O fenótipo é resultado do patrimônio genético que o animal possui, somado aos efeitos de meio ambiente, existindo ainda a interação entre os efeitos de genótipo e de meio ambiente

A interação genótipo x ambiente aborda o conceito de que o melhor genótipo em um ambiente poderia não o ser em outro (FALCONER & MACKAY, 1996). Assim, animais de raça especializada, com alta produção de leite em clima temperado, podem não apresentar o mesmo desempenho produtivo em países de clima tropical.

Para que os objetivos estabelecidos pelo produtor sejam alcançados, é necessário obter informações confiáveis dos animais que serão selecionados para acasalamento. A exatidão dessas avaliações e os acertos das decisões de acasalamento são, entre outros fatores, determinantes das taxas de progresso genético que indicam a eficácia dos programas de seleção. As técnicas ótimas de avaliação envolvem, simultaneamente, a predição de valores genéticos e a estimação de componentes de (co)variância, sob modelos estatísticos em nível de indivíduos (NEVES, 2007).

RNAs têm sido utilizadas nos últimos anos em larga escala no setor agropecuário. MENDES et al (2009) estudaram redes neurais artificiais do tipo perceptron multicamadas nas predições univariada e bivariada do peso ao sobre ano em bovinos de corte. O trabalho objetivou a predição do peso ao sobre ano da progênie (P550BN) de fêmeas da raça Nelore, por meio de análises univariada e bivariada. Na análise univariada utilizou-se a característica peso ao desmame da vaca (P205N) como variável de entrada, sendo que na análise bivariada as características de entrada foram peso ao desmame da vaca (P205N) e peso ao ano da vaca (P365N). A rede neural artificial

(RNA) realizou a predição do fenótipo peso ao sobre ano da progênie (P550BN), com erro quadrático médio (EQM) igual a 0,0001 na interação de ambas as análises. A correlação linear, entre os valores reais (VR) e valores preditos (VPR) pela rede neural artificial (RNA) foi 0,8657 para a análise univariada e 0,7016 para a análise bivariada.

BRENNEKE (2007) utilizou a RNA para a predição de proteína bruta e suas frações usando base de dados que incluiu as variáveis tamanho, cor das folhas, média da altura pré-pastejo, latitude e longitude, perfilho novos e perfilho remanescentes. Neste trabalho, a partir de dados externos da planta a pesquisadora demonstrou que a RNA conseguiu estabelecer relação entre estas variáveis externas e seu conteúdo de proteína bruta.

Em estudos visando a previsão de mastite bovina, as RNAs foram usadas como sistema de diagnóstico precoce baseado em dados de monitoramento diário coletados durante as ordenhas e armazenados em uma base de conhecimento para assim detectar com antecedência o surgimento da mastite bovina nos animais em estudo (CAVERO et al., 2008; HEALD et al., 2000).

NEVES (2007) estudou RNAs do tipo perceptron multicamadas para predição dos valores genéticos em características relacionadas à produção em gado de leite e comparou os resultados dessa metodologia com os obtidos pelo melhor preditor linear não viesado (*Best Linear Unbiased Predictor* – BLUP), gerados por análises feitas por meio do programa

MTDFREML (*Multiple Trait Derivative Free Restricted Maximum Likelihood*). Para os dados de entrada da rede foram utilizados registros, coletados entre 1998 e 2002, referentes à produção de leite total, de gordura total, período de lactação e idade da vaca ao parto de 2.500 vacas da raça Pardo-Suíça. Por se tratar de uma rede com aprendizado supervisionado foram utilizados como saída desejada os valores genéticos preditos pelo BLUP de cada uma dessas características. A comparação entre os resultados da RNA com os do BLUP apresentou uma correlação positiva moderada, indicando o potencial das RNAs na avaliação genética em bovinos de leite.

MEIRELLES (2005) estudou RNAs na avaliação e predição de valores genéticos para ganho de peso em bovinos de corte de raças zebuínas brasileiras. As características estudadas foram peso ao nascimento, peso ao desmame, peso ao sobreano e ganho de peso do desmame ao sobreano (GP345). Segundo a pesquisadora, os resultados apresentados indicaram o potencial das RNAs na avaliação genética, grande melhoria na velocidade de obtenção dos dados e menor custo computacional necessário, fatos que justificam a continuidade dessa linha de pesquisa. BERTAZZO (2006) pesquisou RNAs na predição do mérito genético de bovinos de corte, justificando seu uso em programas de melhoramento genético (Tabela 1).

Embora a rede neural artificial (RNA) tenha provado ser eficiente ferramenta de classificação e de regressão, a compreensão das relações estabelecidas entre a entrada e as variáveis de saída é um problema. O reconhecimento de

padrões é definido como o processo de identificação de estrutura de dados válidos, potencialmente úteis e compreensíveis (BRADLEY et al, 1998).

Tabela 1. Principais estudos nacionais com uso de RNAs para predição do valor genético

<i>Autores</i>	<i>Atuação</i>
MEIRELLES (2005)	Bovinos de corte
NEVES (2006)	Bovinos de leite
BERTAZZO (2007)	Bovinos de corte

O presente estudo tem por objetivo avaliar o uso das RNAs na predição de valores genéticos do peso aos 205 (P205) dias de idade de animais da raça Tabapuã e avaliar alterações dos resultados classificatórios desses animais quando se faz uso de RNAs treinadas por meio do algoritmo de treinamento *LM - Levenberg Marquardt*.

## 2 - MATERIAL E MÉTODOS

As informações utilizadas neste estudo são de animais Tabapuã, nascidos entre 1976 e 1995, provenientes de 152 fazendas dos diversos estados brasileiros, agrupados em quatro regiões distintas [Sul (R1), Sudeste (R2), Centro-Oeste (R3) e Norte/Nordeste (R4)] incluídos no Controle de Desenvolvimento Ponderal da Associação Brasileira de Criadores de Zebu (ABCZ) e cedidos pela Embrapa Gado de Corte, localizada em Campo Grande, MS.

A característica estudada foi o peso do animal aos 205 (P205) dias de idade. Os pesos foram padronizados por interpolação, e o intervalo considerado entre as pesagens foi, no máximo, de 40 dias, de acordo com o sistema de avaliação genética adotado pelo Controle de Desenvolvimento Ponderal da ABCZ.

Foram utilizados como base de dados inicial no presente estudo 19240 registros de animais da Raça Tabapuã contidos no banco de dados, após restrições aplicadas aos mesmos.

Foram consideradas duas estações de nascimento, de acordo com as épocas das águas e das secas das regiões. Foram também definidos os "touro de conexão", isto é, touros com número mínimo de duas progênes em cada uma de três ou mais regiões. Em seguida, os grupos contemporâneos (GC) foram formados por ano de nascimento, fazenda, estação de nascimento e sexo. As regiões que apresentavam número de GC menor do que três foram eliminadas.

Para obtenção dos valores genéticos por meio do melhor preditor linear não viesado (*Best Linear Unbiased Predictor - BLUP*), utilizou-se o software WOMBAT (MEYER, 2006), alimentado com registros referentes ao identificador do animal (ANI), PAI, MAE, GC (grupo contemporâneo), identificador da região (REG), idade da mãe ao parto (IDMP) e peso aos 205 dias (P205). Considerou-se como efeito fixo as variáveis GC e REG, a variável IDMP foi considerada como covariável no modelo. A Tabela 2 contém amostras desses dados para melhor compreensão.

Tabela 2. Amostra de registros dos animais da raça Tabapuã contidos no banco de dados EMBRAPA GADO DE CORTE (MS)

<i>ID_ANI</i>	<i>ID_PAI</i>	<i>ID_MAE</i>	<i>ID_GC</i>	<i>ID_REG</i>	<i>IDMP(dias)</i>	<i>PNASC(kg)</i>	<i>P205(kg)</i>
28643	132	5929	9601222	1	3956	24	143
21244	109	12070	9100521	2	2509	34	191

GC composto por ano de nascimento, fazenda, estação de nascimento e sexo.

A Tabela 3 ilustra os valores da média geral e coeficiente de variação, para as variáveis peso ao nascimento, peso aos 205 dias de idade e idade da mãe ao parto dos dados utilizados neste estudo

Tabela 3 – Valores da média e coeficiente de variação (CV%) de peso aos 205 dias (P205), peso ao nascimento (PNASC) e idade da mãe ao parto (IDMP)

<i>Variável</i>	<i>Média</i>	<i>CV%</i>
PNASC (kg)	31,17	9,1
P205 (kg)	171,86	14,6
IDMP (dias)	2801	20,3

Neste trabalho utilizou-se a rede neural artificial (RNA) do tipo perceptron multicamadas (MultiLayer Perceptron - MLP) e algoritmo de treinamento supervisionado *LM - Levenberg Marquardt* para reconhecimento de padrões, comparação da saída RNA com os parâmetros obtidos no BLUP e futura classificação dos animais de acordo com a saída BLUP e saída RNA. Segundo MEIRELLES (2005), na fase de treinamento, a rede aproxima a saída ao valor desejado e executa o algoritmo de treinamento escolhido.

Para desenvolvimento da RNA, utilizou-se a toolbox ou biblioteca NEURAL NETWORK TOOLBOX do MATLAB (2006). O algoritmo utilizado no estudo *LM - Levenberg Marquardt* é definido

nesse software como TRAINLM, como aperfeiçoamento do tradicional algoritmo *Backpropagation*.

O algoritmo LM utiliza o erro quadrático médio ou MSE (*mean squared error*) para corrigir os pesos da rede quando sua saída for diferente da saída desejada, denominadas pelos *target's*. (MEIRELLES,2005)

Como entradas da RNA foram inseridas as variáveis *idade da mãe ao parto* (IDMP), *estação (seca ou águas codificadas como “1” ou “2” respectivamente)*, *região, peso aos 205 dias de idade* (P205) e o valor genético obtido pelo BLUP sendo “target” ou objetivo da RNA. O estudo em questão é dependente das saídas dos valores genéticos obtidos pelo BLUP, após utilização do software WOMBAT (MEYER, 2006).

Segundo o mesmo procedimento adotado por NEVES (2007) em seu estudo, com o objetivo tornar o treinamento da rede mais eficiente, as entradas da RNA foram normalizadas para o intervalo [-1,1] pela faixa de variação listada abaixo:

$$x_{norm} = \frac{(x - x_{min})}{(x_{max} - x_{min})}$$

em que,  $x_{norm}$  é o valor normalizado correspondente ao valor original  $x$ , e  $x_{min}$  e  $x_{max}$  são o valor mínimo e máximo entre todos os dados. O trecho abaixo descreve código fonte desenvolvido em linguagem MATLAB para tal processo de normalização.

```
% -----
% TRECHO DO PROGRAMA MATLAB
% Faz a normalizacao dos dados para valores
entre -1 e 1.
[DTn,minDTn,maxDTn] = premmx(DT);
[TTn,minTTn,maxTTn] = premmx(TT);
% Transformacao dos dados usando min e max
precalculado
[DVn] = trammmx(DV,minDTn,maxDTn);
[TVn] = trammmx(TV,minTTn,maxTTn);
% -----
```

A base de dados foi dividida em dois conjuntos de dados, destinados ao treinamento e validação. A proporção

adotada foi de 75% e 25% para treinamento e validação, respectivamente. O primeiro conjunto destina-se ao ajuste de pesos sinápticos e o segundo à fase de avaliação e desempenho da rede (Tabela 4).

A camada de entrada é composta por quatro neurônios, sendo estes formados pelos componentes *idade da mãe ao parto* (IDMP), *estação (seca ou águas codificadas como “1” ou “2”* respectivamente), *região*, *peso aos 205 dias de idade* (P205). A camada oculta é composta por N neurônios, de acordo com os testes de desempenho realizados e demonstrados na Tabela 5. A camada de saída é composta pelo valor genético obtido pela RNA (VG\_RNA) após processo de treinamento e posterior validação (Figura 1).

Tabela 4. Divisão do conjunto de dados inicial

CATEGORIA	No de dados	%
TREINAMENTO	14430	75
TESTE	4810	25

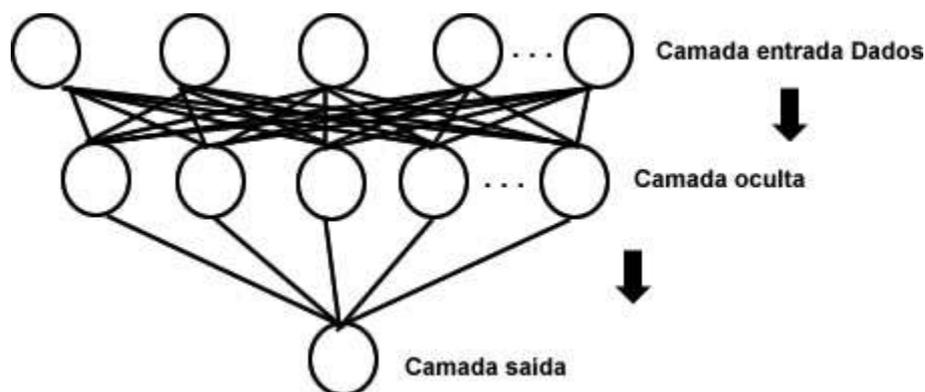


Figura 1. Conexões neurais em uma RNA multicamada.

Segundo NEVES (2007), o número de camadas intermediárias e de neurônios presentes nessa camada deve ser projetado mantendo o compromisso com a convergência para uma solução, sem, no entanto, prejudicar a habilidade de generalização da rede neural. Dessa forma, foram desenvolvidas e comparadas diferentes propostas de arquiteturas de redes MLP, que se diferenciaram pelo número de neurônios

presentes na camada intermediária e taxa de aprendizado. A taxa de aprendizagem ( $\eta$ ) influencia as mudanças nos pesos entre as conexões dos neurônios artificiais. Taxas de aprendizagem muito pequenas (próximas de zero) implicam em tempo de treinamento longo; contudo, taxas de aprendizagem elevadas podem provocar oscilações em torno da solução.

Tabela 5. Propostas de utilização da RNA para predição do valor genético VG\_P205\_RNA

<i>Proposta</i>	<i>Camadas ocultas</i>	<i>Neurônios ocultos</i>	<i>Taxa aprendizado</i>	<i>Erro meta</i>	<i>Iterações (Épocas)</i>
A	1	30	0,01	0,01	1000
B	1	40	0,01	0,01	500
C	1	45	0,01	0,01	100
D	1	45	0,008	0,01	100
E	1	45	0,1	0,01	100

Todas as arquiteturas tiveram como método de treinamento o algoritmo *Levenberg Marquardt* (TRAINLM – MATLAB, 2006) e o conjunto de pesos inicializado aleatoriamente dentro do intervalo [-1,1]. Como parâmetros de treino têm-se: a função de ativação, a quantidade máxima de épocas ou iterações, taxa de aprendizagem e valor da tolerância ao erro. As funções de ativações utilizadas na camada intermediária e de saída foram tangente hiperbólica (TANSIG) e linear (PURELIN), respectivamente.

análise adicional foi realizada com esta arquitetura para avaliar o poder da RNA na predição dos valores genéticos de animais que fossem inseridos posteriormente ao último ano adotado como limite de treinamento da rede neural. Foi realizado treinamento da rede neural com informações de animais contidas na base de dados até o ano de 1990 e para o processo de validação foi utilizada informações de animais pertencentes aos anos 1991 a 1995, divididos em duas classes (1991-1993 e 1994-1995).

Para o treinamento da RNA, o número de épocas foi escolhido como critério de parada, de acordo com a Tabela 6, assim como os valores de neurônios ocultos e taxa de aprendizado.

Após obtenção da melhor arquitetura de RNA para predição do valor genético,

### 3. RESULTADOS

O valor de erro pode interromper prematuramente o processo de aprendizagem ou fazer com que a rede perca sua capacidade de generalização (HAYKIN, 2001).

De acordo com NEVES (2007), o desempenho de cada arquitetura foi observado com base no erro quadrático médio (MSE ou EQM), obtido pela média dos quadrados das diferenças

entre a saída real e a saída prevista pela rede, assim demonstrado a seguir:

$$EQM = \frac{1}{N} \sum_{i=1}^N (VG_{BLUP} - VG_{RNA})^2$$

Tabela 6. Valores do erro quadrático médio de acordo com as propostas estabelecidas

<i>Proposta</i>	<i>Camadas ocultas</i>	<i>Neurônios ocultos</i>	<i>Taxa aprendizado</i>	<i>Erro meta</i>	<i>Iterações (Épocas)</i>	<i>EQM Validação</i>
A	1	30	0,01	0,01	1000	0.0111717
B	1	40	0,01	0,01	500	0.0110312
C	1	45	0,01	0,01	100	0.0111102
D	1	45	0,008	0,01	100	0.0110078
E	1	45	0,1	0,01	100	0.0110756

Em análise da Tabela 6, a proposta D, com o menor EQM de teste, é constituída por quarenta e cinco neurônios na camada intermediária e taxa de aprendizado igual a 0,008.

estabelecida pela proposta D (melhor desempenho da RNA), por intermédio da estimativa do coeficiente de correlação de Pearson (r), que foi 0,8503.

Entretanto, como observado, o fato relevante para o desempenho da RNA foi o número de camadas ocultas seguido pela taxa de aprendizado. O número de iterações não interferiu de forma significativa no valor EQM, sendo este estabilizado em todas as propostas por volta da décima quinta época, conforme ilustra a figura 2.

Este alto coeficiente de correlação entre os valores genéticos preditos pela RNA e os obtidos pelo BLUP, indica que as variáveis envolvidas no estudo estão muito associadas e têm magnitude de variação no mesmo sentido. Conforme figura 3, a RNA reconheceu o padrão de comportamento dos dados e conseguiu prever de forma satisfatória o valor genético para peso aos 205 dias, para os dados estudados. Os resultados assemelham-se aos obtidos por MEIRELLES (2005) e NEVES (2007), em análise de outras características por meio de RNAs (r = 0,877 e r = 0,65, respectivamente). Caso haja inserção de novos animais para nova avaliação genética, novo treinamento deverá ser efetuado e novos targets deverão ser gerados para construção da RNA, sempre dependentes do BLUP.

Segundo PEREIRA (2004), o termo correlação descreve a associação entre duas variáveis e o grau de relacionamento entre elas é medido pelo coeficiente de correlação, que pode oscilar de zero a um, de forma positiva ou negativa.

Os valores genéticos oriundos do BLUP (VG\_P205\_BLUP) foram comparados aos obtidos pela RNA (VG\_P205\_RNA)

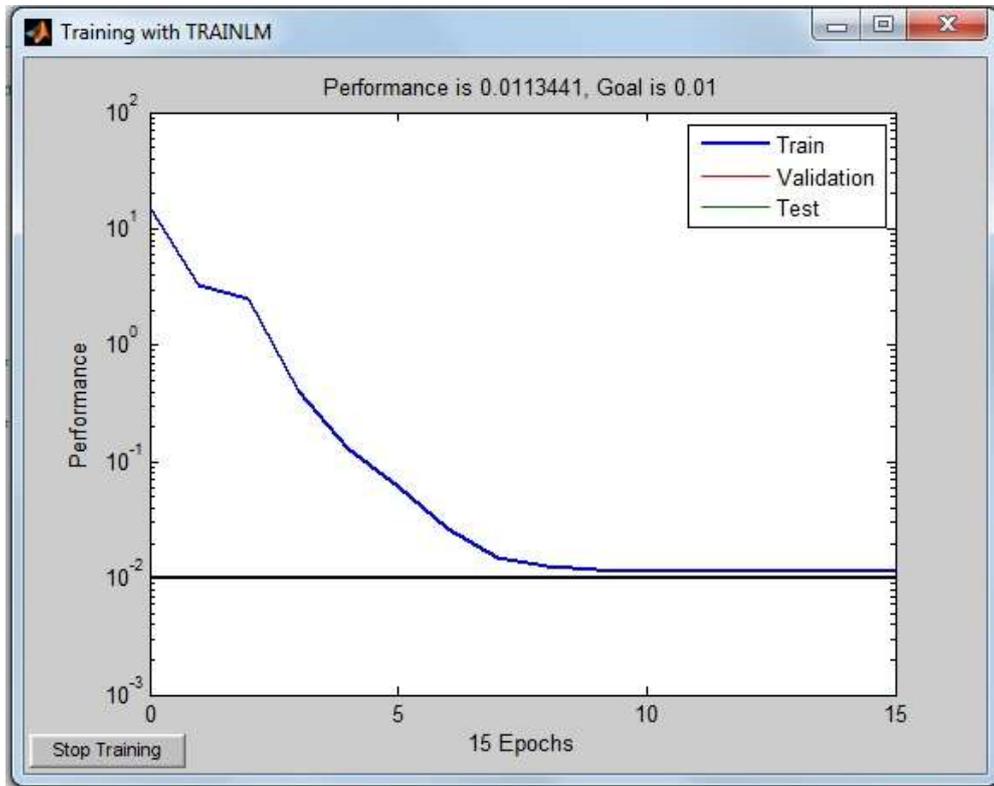


Figura 2. Desempenho obtido pela RNA de acordo com o número de épocas.

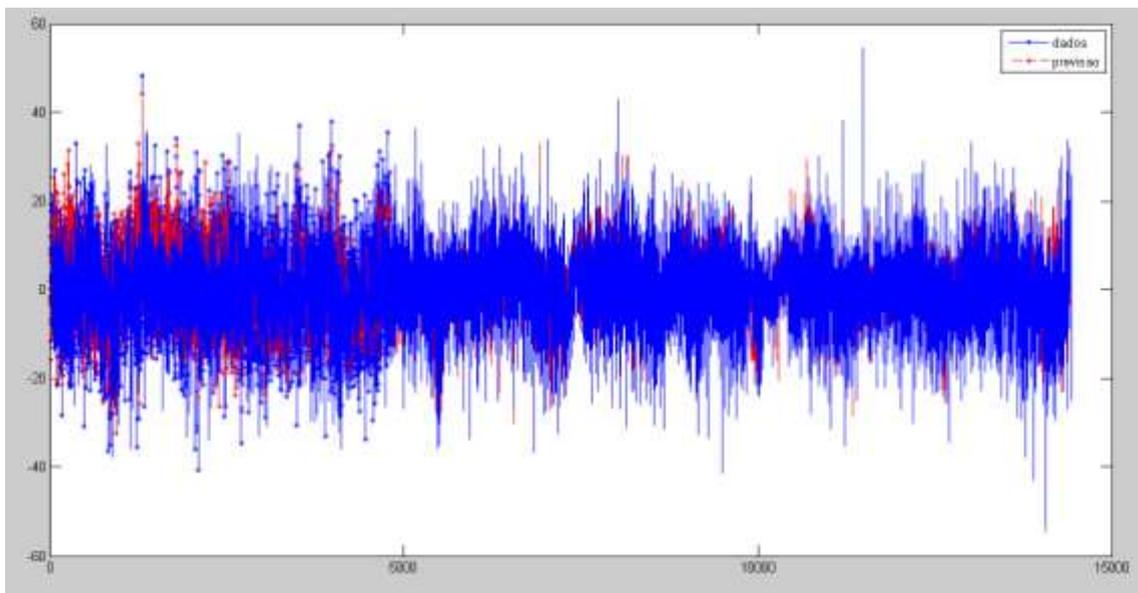


Figura 3. Predição dos valores genéticos para peso aos 205 dias por intermédio de RNA.

Após ordenar os valores genéticos para característica peso aos 205 dias (VG\_P205\_RNA), oriundos da RNA, percebeu-se variação na classificação dos animais em comparação à classificação obtida pelo BLUP, indicando riscos no uso de RNAs para avaliação genética para esta característica (Tabela 7A).

Após classificação baseada em valores genéticos de peso aos 205 dias, preditos pelo BLUP e RNA, foram calculadas as percentagens de coincidência de

classificação entre os trezentos primeiros colocados das duas classificações (Tabela 7B). Os primeiros cem classificados por meio da RNA não coincidiram em 66% com a classificação obtida por meio do BLUP. As classificações posteriores apresentaram menor valor de coincidência (30 e 22%, respectivamente) do que apresentada na primeira faixa pelo fato dos valores genéticos dos animais pertencentes a essa classe apresentarem menores desvios.

Tabela 7A. Classificação baseada em valores genéticos de peso aos 205 dias preditos pelo BLUP (P205-BLUP) e RNA (P205-RNA) para animais da raça Tabapuã

<i>Animal – Código Interno</i>	<i>Classificação BLUP</i>	<i>Classificação RNA</i>
<b>001</b>	<b>1<sup>o</sup></b>	<b>1<sup>o</sup></b>
<b>002</b>	<b>2<sup>o</sup></b>	<b>2<sup>o</sup></b>
<b>003</b>	<b>-</b>	<b>3<sup>o</sup></b>
<b>004</b>	<b>7<sup>o</sup></b>	<b>4<sup>o</sup></b>
<b>005</b>	<b>-</b>	<b>5<sup>o</sup></b>
<b>006</b>	<b>8<sup>o</sup></b>	<b>6<sup>o</sup></b>

Tabela 7B. Percentagens de valores classificatórios coincidentes e não coincidentes entre classificação obtida por meio do BLUP e RNA

<i>Classificação</i>	<i>% de coincidência</i>	<i>% de não coincidência</i>
<b>1 – 100</b>	<b>34%</b>	<b>66%</b>
<b>101 – 200</b>	<b>30%</b>	<b>70%</b>
<b>201 – 300</b>	<b>22%</b>	<b>78%</b>

A Tabela 8 mostra as correlações obtidas entre saídas da RNA e valor BLUP original para animais que foram inseridos na análise após processo de treinamento ter sido finalizado no ano de 1990. O primeiro grupo foi representado

por animais inseridos no processo de avaliação a partir do ano de 1991 até 1993, apresentou menor correlação com os dados de saída BLUP ( $r = 0,74$ ) do que o valor encontrado na análise anterior, tendo base de dados formada

por animais com informações nos anos de 1976 a 1995 ( $r = 0,8503$ ). Para o processo de validação da RNA, por meio de informações do segundo grupo de animais inseridos na base de dados nos anos 1994 e 1995, a queda de associação foi ainda maior ( $r = 0,68$ ).

Em análise complementar, foi utilizada mesma base de dados para execução do algoritmo MOBJ em comparação aos resultados obtidos pelo algoritmo *Levenberg Marquardt* (LM) (Tabela 9).

Tabela 8. Correlação de Pearson ( $r$ ) entre saídas BLUP e RNA, para animais inseridos no processo de validação após processo de treinamento ter sido finalizado.

Ano	Correlação Pearson ( $r$ ) BLUP – RNA
1991 – 1993	0,74
1994-1995	0,68

Tabela 9. Acuracias obtidas pelo algoritmo MOBJ

Proposta	AcuraciaValidação
A	99.4%
B	99.7%
C	99.5%
D	99.8%
E	99.7%

Conforme ilustra a Tabela 9, os resultados obtidos pelo algoritmo MOBJ indicaram semelhança funcional quando comparados à saída do algoritmo LM (Tabela 6).

### 3. CONCLUSÕES

- As RNAs demonstram grande versatilidade referente ao desempenho operacional e ao tempo de processamento, com grande vantagem em relação a outras metodologias de predição.
- Para mesmo banco de dados, sem inserção de novos animais, a RNA é eficaz no reconhecimento de padrões e predição do valor genético para a característica peso aos 205 dias, utilizando como target o valor genético oriundo da metodologia BLUP.
- Não é recomendado o uso de RNAs em avaliações genéticas quando se deseja inserir no futuro novos animais que não estejam contidos na base de dados treinados. A RNA é dependente da metodologia BLUP por não conter informações da matriz de relacionamento entre animais.

### 5.REFERENCIAS BIBLIOGRÁFICAS

BERTAZZO, R. P. *Utilização de redes neurais artificiais na predição do mérito genético de bovinos de corte*. Tese (Doutorado) - Faculdade de Zootecnia e Engenharia de Alimentos, Universidade Federal de Minas Gerais, Pirassununga, SP, 2006.

BRADLEY, P. S., FAYYAD, U. M.; MANGASARIAN, O. L. Mathematical programming for datamining: formulation and challenges. *NFORMS Journal on Computing*, v.3, n.11, p.217-238, 1998.

- BRENNECKE, K. *Fracionamento de carboidratos e proteínas e a predição da proteína bruta e suas frações e das fibras em detergentes neutro e ácido de Brachiaria brizantha cv. Mandaru por uma rede neural artificial*. 2007. 138f. Tese (Mestrado em Zootecnia) – Faculdade de Zootecnia e Engenharia de Alimentos, Universidade de São Paulo, Pirassununga, SP.
- CAVERO, D.; TOLLE, K. H.; HENZE, C. *et al.* Mastitis detection in dairy cows by application of neural networks. *Livestock Science*, v.114, n.2-3, p.280-286, 2008.
- FILHO, E. K. *Melhoramento Genético Animal no Brasil: fundamentos, história e importância*. Campo Grande: Embrapa, 1999.
- FALCONER, D. S.; MACKAY, T. F. C. *Introduction to quantitative genetics*. 4.ed. Harlow: Longman, 1996. 464 p.
- HAYKIN, S. *Redes neurais: princípios e prática*. Porto Alegre, RS: Bookman, 2001.
- HEALD, C. W.; KIM, T.; SISCHO, W. M. *et al.* A computerized mastitis decision aid using farm – based records: an artificial neural network approach. *Journal of Dairy Science*, v.83, n.4, p.711-720, 2000.
- KEPLER FILHO, E. *Melhoramento genético animal no Brasil: fundamentos, história e importância*. Campo Grande: Embrapa, 1999.
- MATLAB guide for personal computers. Version 7.3. Local: MATLAB Institute, 2006.
- MEIRELLES, F. D. P. *Modelo computacional de um rebanho bovino de corte virtual utilizando simulação Monte Carlo e redes neurais artificiais*. 2005. 104f. Tese (Doutorado em Zootecnia) – Faculdade de Zootecnia e Engenharia de Alimentos, Universidade de São Paulo, Pirassununga, SP.
- MENDES, R. A.; BERTAZZO, R. P.; SILVA, D. C. *et al.* Sistemas especialistas na predição do fenótipo peso ao sobreano em bovinos da raça Brahman. In: SIMPÓSIO INTERNACIONAL DE INICIAÇÃO CIENTÍFICA DA USP (SIICUSP), 2009, Pirassununga SP. *Anais...* Pirassununga,: USP, 2009.
- MEYER, K. WOMBAT: digging deep for quantitative genetic analyses by restricted maximum likelihood. In: WORLD CONGRESS ON GENETIC APPLIED TO LIVESTOCK PRODUCTION, 8., 2006, Belo Horizonte. *Proceedings...* Belo Horizonte: Instituto Prociência, 2006. CD-ROM.
- NEVES, A.R.M. *Aplicação de redes neurais artificiais na predição de valores genéticos em bovinos de leite da Raça Pardo-Suiça*. Belém: Universidade Federal do Pará, Instituto de Tecnologia, 2007.
- PEREIRA, J. C. *Melhoramento genético aplicado à produção animal*. Belo Horizonte, FEP-MVZ, 2004. 610p.

## CAPITULO 2

### **PREDIÇÕES DA PRODUÇÃO DE LEITE NO SEGUNDO E TERCEIRO PARTOS POR INTERMÉDIO DE INFORMAÇÕES DOS PARTOS ANTERIORES UTILIZANDO REDES NEURAIAS ARTIFICIAIS E COMPARAÇÃO COM A CAPACIDADE PROVÁVEL DE PRODUÇÃO.**

#### **RESUMO**

Sistemas baseados em redes neurais artificiais (RNAs) podem ser utilizados como ferramenta para tomada de decisões em propriedades rurais como também em diversos setores industriais agrícolas. Pesquisadores e produtores podem utilizar tal metodologia para prever produções futuras em vacas leiteiras de diferentes graus de sangue, baseados no reconhecimento de padrão e relações não lineares descobertos pelas RNAs. Dados de 2667 animais foram utilizados para treinamento da RNA por meio do algoritmo LM - Levenberg Marquardt, para predição da segunda lactação por intermédio de informações da primeira lactação. Em segunda análise, dados de 1636 animais que apresentavam informações da primeira, segunda e terceira lactação foram utilizados para predição da terceira lactação por meio de RNAs. A capacidade provável de produção (CPP) foi calculada e comparada com os dados originais da produção de leite em 305 dias de lactação e os resultados da RNA. Os resultados obtidos pela RNA apresentaram forte associação com os dados de produção de leite observados e os preditos por meio da CPP, sendo esta última considerada forma mais simplificada para extração desta informação.

**Palavras-chave:** Produção de leite, Redes Neurais Artificiais, Algoritmo Levenberg Marquardt

## ABSTRACT

Systems based on artificial neural networks (ANNs) can be used as a tool for decision in farms and in many agricultural industries. Researchers and producers may use the methodology to predict future production of dairy cows of different genetic composition, based on pattern recognition and nonlinear relationships using the ANNs procedure. Data from 2667 animals were used in ANNs procedure and LM algorithm - Levenberg Marquardt for training purpose to predict second lactation production based on information of the first lactation. In the second analysis first, second and third lactation records from 1636 animals were used to predict the third lactation using ANNs procedure. The milk producing ability (MPA) was calculated and compared with the original data for milk production in 305 days of lactation and to the results obtained from the ANNs procedure. The results obtained from ANNs procedure were strongly associated with the observed milk production data and the milk producing ability (MPA), the last one being a simplified form for obtaining this information.

**Keywords:** Milk production, Levenberg Marquardt algorithm

## 1- INTRODUÇÃO

Algoritmos baseados em redes neurais artificiais (RNAs) processam informações utilizando grande número de neurônios e reconhecem padrões por mecanismos de aprendizagem. O conhecimento é armazenado dentro de conexões entre os neurônios (SAMARASINGHE, 2006). Para diferentes efeitos, várias RNAs podem ser desenvolvidas com base em diferenças no arranjo das camadas, interconexão de elementos, e nos métodos de aprendizagem.

Tais neurônios processam os dados de entrada vinculados ao objetivo de saída da RNA ou *target* da rede. A RNA busca reconhecimento de padrões entre dados de entrada, e saída desejada, extraindo relações não lineares, para que após o processo de treinamento, gere automaticamente dados semelhantes aos dados da saída apresentados como modelo ou *target* desejado. A

capacidade de aprender por meio de exemplos e de generalizar a informação aprendida é o atrativo principal da solução de problemas por meio de RNAs. A generalização, que está associada à capacidade de aprendizado por meio de restrito conjunto de dados e em segundo momento gerar informações coerentes para dados não conhecidos é a demonstração de que as RNAs não mapeiam apenas relações de entrada e saída (BRAGA et al, 2007).

Divergências em desempenho de RNAs são reportadas na área agrária em consequência das diferentes possibilidades de estruturação da RNA (SABLANI et al., 1995; YANG et al., 2000). RNAs têm sido utilizadas para estimação de contaminação por aflatoxina em lavouras de amendoim antes da colheita, desenvolvimento de plantas, produtividade de milho e parâmetros de lucratividade em indústrias de laticínios. (HOSSEINIA, 2007).

Na década de 90, RNAs foram utilizadas para avaliação e classificação de animais baseados na produção de leite em 305 dias de lactação (LACROIX et al, 1995 ; SALEHI et al, 1998).

GRZESIAK et al(2006) compararam a capacidade preditiva da produção de leite em 305 dias por dois métodos diferentes, função de Wood e RNAs. A predição via RNAs apresentou maior acurácia do que em modelos que utilizaram função de Wood.

NEVES (2007) utilizou registros referentes à produção de leite total, gordura total, período de lactação e idade da vaca ao parto de 2500 vacas Pardo-Suiça para predição de valores genéticos por meio de RNAs. A comparação entre os resultados da RNA com os do BLUP apresentou uma correlação positiva (0,65), indicando o potencial das RNAs na avaliação genética em bovinos de leite.

SUN et al (2009) concluíram que a RNA é ferramenta eficaz para detecção antecipada de mastite em bovinos de leite por meio de informações pré-processadas de dados relacionados à condutividade elétrica e produção de leite.

Os pesquisadores NAZARIO et al (2007) avaliaram a caracterização de leite bovino por meio de ultrasonografia e redes neurais artificiais, para utilização em indústrias de laticínios. Observou-se que as redes neurais possibilitaram grande poder de classificação e isso foi demonstrado nos experimentos de testes com amostras desconhecidas pelas redes projetadas. Para o problema de adição de água obteve-se um poder de

classificação de 96,3%, enquanto que para a obtenção do teor de gordura, 95,5% de classificação correta.

NJUBI et al (2009) pesquisaram registros provenientes do dia do controle de animais Holstein-Friesian para predição da produção de leite no mês seguinte e em 305 dias de lactação por meio de RNAs. Os resultados foram comparados com os observados fazendo-se uso de um modelo de regressão linear múltipla, e com índice de acurácia na estimativa dos parâmetros de 75%.

A utilização de médias de produções do mesmo animal, em diferentes etapas de sua vida produtiva, contribui para aumentar a precisão na identificação dos melhores genótipos da população. Avaliação e seleção de fêmeas dentro de um rebanho estão relacionadas à maior ou menor capacidade de julgar os desempenhos dos animais naquele rebanho. Diferenças ambientes temporárias, que ocorrem entre os anos, contribuem para gerar variações no desempenho dos animais e conseqüentemente, dificuldades na identificação dos melhores genótipos. Altas produções dentro de um rebanho resultam de genótipos superiores e ambientes favoráveis. A separação dos efeitos ambientes temporários daqueles permanentes é a base segura para a seleção de fêmeas dentro de um rebanho. O uso da média de  $n$  observações repetidas é um dos modos mais eficientes de se controlarem erros e confusões que de outra maneira poderiam resultar dos efeitos temporários ambientes. Tendo-se a média de  $n$  produções pode-se obter a sua capacidade provável de produção (CPP), que inclui os efeitos permanentes

de meio, além do valor genético do animal (PEREIRA, 2004).

O presente estudo tem por objetivo avaliar o uso das RNAs na predição da segunda lactação por meio de informações da primeira lactação, de animais da Raça Gir Leiteiro e cruzamentos com animais da Raça Holandesa, da terceira lactação quando são disponíveis dados da primeira e segunda lactações. O algoritmo de treinamento *LM - Levenberg Marquardt* é utilizado nesse estudo para tentativa de reconhecimento de padrões entre os dados inseridos na rede neural artificial (RNA). Os resultados obtidos pela RNA serão comparados com a capacidade provável de produção calculada à priori e os dados reais de produção.

## 2 - MATERIAL E MÉTODOS

Um total de 145.079 registros da produção de leite no dia do controle (PLDC) de vacas Gir leiteiro e animais cruzados Holandês-Gir, paridas entre os anos de 1990 e 2006, com idade inicial ao parto de 24 meses, provenientes de rebanhos dos diversos estados brasileiros foi utilizado neste estudo. As informações foram obtidas do banco de dados do Programa Nacional de Melhoramento do Gir Leiteiro (PNMGL) e Associação Brasileira de Criadores de Zebu (ABCZ), cedidos pela Empresa Brasileira de Pesquisa Agropecuária (EMBRAPA – Gado de Leite), localizada em Juiz de Fora, MG.

Após filtro de dados, detectaram-se 57.246 registros de animais em primeira lactação e 30.621 registros na segunda lactação. Em busca de animais que

apresentassem ao mesmo tempo registros de primeira e segunda lactação, 3556 animais foram selecionados no banco de dados. Posteriormente, foram também selecionados 1636 animais que possuíam registros em comum no conjunto de dados da primeira, segunda e terceira lactações.

Os animais em estudo apresentavam a seguinte composição genética (Tabela 1):

Tabela 1. Diferentes composições genéticas dos animais em estudo

<i>Bases Genéticas</i>	<i>Fração GIR x HOLANDES</i>
1	0,0 HOL 1,0 GIR
2	0,5 GIR 0,5 HOL
3	0,75 GIR 0,25 HOL
4	0,75 HOL 0,25 GIR
5	0,875 GIR 0,125 HOL
6	0,625 GIR 0,375 HOL
7	0,625 HOL 0,375 GIR
8	0,875 HOL 0,125 GIR
9	0,3125 HOL 0,6875 GIR
10	0,3125 GIR 0,6875 HOL

A característica estudada foi produção de leite aos 305 dias (PL305), com variação de informações do primeiro ao terceiro parto.

O controle leiteiro, geralmente realizado em intervalos de 30 dias, é uma prova zootécnica que consiste na mensuração e correspondente registro da produção individual da vaca em um período de 24 horas. Os registros do controle leiteiro constituem-se então na única fonte de informação das avaliações genéticas de bovinos para a produção de leite. Tradicionalmente, tais registros são utilizados para a estimação da produção acumulada até 305 dias, medida esta

utilizada nas avaliações genéticas (JUNQUEIRA, 2009).

Foram utilizados os controles entre o 5º e o 305º dia da lactação e estabelecidas as seguintes condições para a inclusão de animais no estudo: primeiro controle leiteiro realizado antes de 45 dias após o parto; intervalo de controles leiteiros entre 15 e 45 dias; número de controles leiteiros maior do que quatro. Foi gerado, adicionalmente às informações de cada animal, o parâmetro informativo de grupo contemporâneo, caracterizado por animais que possuam semelhança em rebanho-ano-época do controle (RAE) com no mínimo três vacas filhas de pelo menos dois touros. Foram eliminados do presente estudo animais com produções maiores ou menores do que três desvios-padrão da média de lactação em cada controle.

A capacidade provável de produção dos animais em cada lactação foi calculada de acordo a seguinte equação:

$$CPP = R + (nr / (1+(n-1)r))(A-R);$$

em que:

CPP = capacidade provável de produção  
R = média do rebanho  
A = média do animal avaliado  
n = número de lactações  
r = repetibilidade da característica, adotou-se a média encontrada na literatura (r = 0,38)

A fração  $(nr/(1+(n-1)r))$  mede a repetibilidade de n produções e fornece a confiabilidade que se pode ter na média de produção do animal como indicadora de sua habilidade produção.

Neste trabalho utilizou-se a rede neural artificial (RNA) do tipo perceptron multicamadas (MultiLayer Perceptron - MLP) e algoritmo de treinamento supervisionado *LM - Levenberg Marquardt* para reconhecimento de padrões.

Para desenvolvimento da RNA, utilizou-se a toolbox ou biblioteca NEURAL NETWORK TOOLBOX do MATLAB (2006). O algoritmo utilizado no estudo *LM - Levenberg Marquardt* é definido nesse software como TRAINLM, sendo um aperfeiçoamento do tradicional algoritmo *backpropagation*.

O algoritmo LM utiliza o erro quadrático médio ou MSE (*Mean Squared Error*) para corrigir os pesos da rede quando sua saída for diferente da saída desejada, denominadas pelos *target's* (MEIRELLES, 2005)

Foram realizadas duas análises distintas no presente estudo. Na primeira análise, dados da primeira lactação serviram como entrada de dados na RNA, tendo como alvo (*target*) de aprendizado a produção de leite em 305 dias da segunda lactação (PL305\_2). Na segunda análise, dados da primeira e segunda lactação formaram a base de dados para entrada da RNA e a produção de leite em 305 dias de lactação (PL305\_3) serviu como *target* ou objetivo da RNA.

## 2.1 Predição da segunda lactação por meio de informações da primeira lactação

Como entradas da RNA foram inseridas as variáveis dias em lactação, produção

de leite em 305 dias no primeiro parto (PL305\_1), idade ao parto, época de parição, média de produção de leite diária, fração de sangue Gir, fração de sangue Holandês, número de ordenhas por dia (2x,3x) e como objetivo de alcance da RNA o parâmetro produção de leite em 305 dias na segunda lactação (PL305\_2). (Figura 1).

A base de dados foi dividida em dois conjuntos de dados, destinados ao treinamento e validação. A proporção adotada foi de 75% e 25% para treinamento e validação,

respectivamente. O primeiro conjunto destina-se ao ajuste de pesos sinápticos e o segundo a fase de avaliação e desempenho da rede (Tabela 2).

A camada de entrada é composta por oito neurônios, sendo estes formados pelos componentes de entrada da rede. A camada oculta é composta por N neurônios. A camada de saída é composta pela estimativa gerada pela RNA para a característica produção de leite no segundo parto (PL305\_2\_RNA) após processo de treinamento e posterior validação (Figura 2).

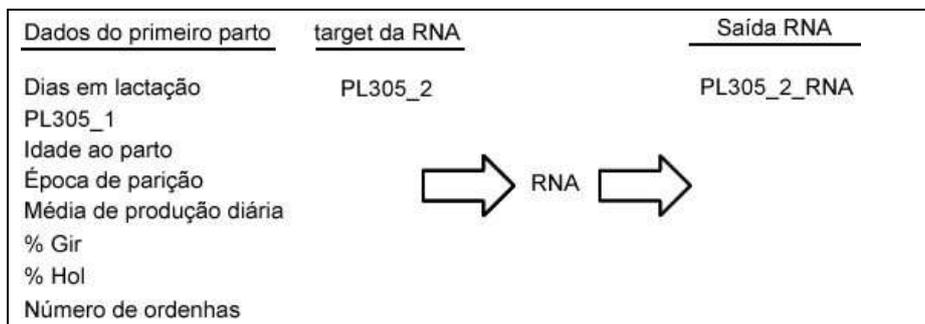


Figura 1. Esquema simplificado da RNA para predição da produção leite no segundo parto.

Tabela 2. Divisão do conjunto de dados inicial

CATEGORIA	No de dados	%
TREINAMENTO	2667	75
TESTE	889	25

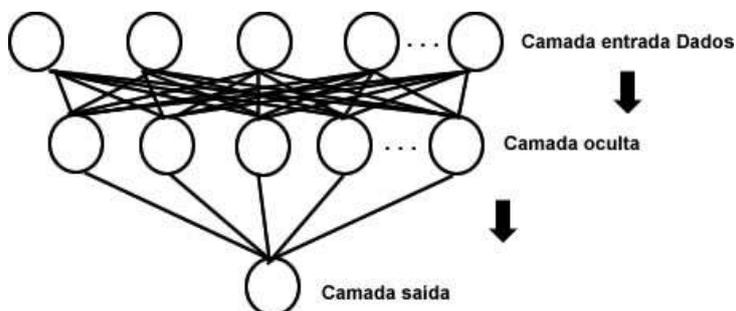


Figura 2. Conexões neurais em uma RNA multicamada.

O número de camadas intermediárias e de neurônios presentes nessa camada deve ser projetado mantendo o compromisso com a convergência para uma solução, sem, no entanto, prejudicar a habilidade de generalização da rede neural (NEVES, 2007). Foram desenvolvidas e comparadas diferentes propostas de arquiteturas de redes MLP, que se diferenciaram pelo número de neurônios presentes na camada intermediária e taxa de aprendizado. A taxa de aprendizagem ( $\eta$ ) influencia as mudanças nos pesos entre as conexões dos neurônios artificiais. Taxas de aprendizagem muito pequenas (próximas de zero) implicam em tempo de treinamento longo; contudo, taxas de aprendizagem elevadas podem provocar oscilações em torno da solução.

Para as duas análises do presente estudo, todas as arquiteturas tiveram como método de treinamento o algoritmo *Levenberg Marquardt* (TRAINLM – MATLAB, 2006) e o conjunto de pesos

inicializado aleatoriamente dentro do intervalo  $[-1,1]$ . Como parâmetros de treino têm-se: a função de ativação, a quantidade máxima de épocas ou iterações, taxa de aprendizagem e valor da tolerância ao erro. As funções de ativações utilizadas na camada intermediária e de saída foram tangente hiperbólica (TANSIG) e linear (PURELIN), respectivamente.

As propostas de utilização da RNA para predição da produção de leite em 305 dias de lactação são apresentadas na Tabela 3. A proposta F foi definida à semelhança da proposta E, entretanto para esse conjunto de dados não houve tratamento e retiradas de exceções na base de dados.

Para o treinamento da RNA, o número de iterações foi escolhido como critério de parada, de acordo com a Tabela 3, assim como os valores de neurônios ocultos e taxa de aprendizado.

Tabela 3. Propostas de utilização da RNA para predição da produção de leite em 305 dias

<i>Proposta</i>	<i>Camadas ocultas</i>	<i>Neurônios ocultos</i>	<i>Taxa aprendizado</i>	<i>Erro meta</i>	<i>Iterações (Épocas)</i>
A	1	30	0.001	0.0001	1000
B	1	30	0.01	0.0001	1000
C	1	35	0.006	0.0001	1000
D	1	60	0.01	0.0001	1000
E	1	65	0.006	0.0001	1000
F	1	65	0.006	0.0001	5000

O valor de erro pode interromper prematuramente o processo de aprendizagem ou fazer com que a rede perca sua capacidade de generalização (HAYKIN, 2001).

O desempenho de cada arquitetura foi avaliado com base no Erro Quadrático Médio (MSE ou EQM), obtido pela média dos quadrados das diferenças entre a saída real e a saída prevista pela rede, assim como apresentado na seguinte equação:

$$EQM = \frac{1}{N} \sum_{I=1}^N (VG_{BLUP} - VG_{RNA})^2$$

Com objetivo de melhor entendimento e visualização das saídas da RNA, parcela dos dados foi escolhida aleatoriamente em cada proposta (A a F) para montagem dos gráficos compostos por dados reais de produção e saída gerada pela RNA para a característica PL305\_2. Os dados completos da proposta F e D, antes e depois de tratamento dos dados, respectivamente, são exibidos nas figuras 8 e 9.

## 2.2 Predição da terceira lactação por meio de informações da primeira e segunda lactação.

As entradas da RNA foram constituídas pelas variáveis dias em lactação (primeiro e segundo partos), produção de

leite aos 305 dias no primeiro parto e segundo partos (PL305\_1 e PL305\_2), idade ao primeiro parto, idade ao segundo parto, época da primeira parição, época da segunda parição, média de produção de leite diária em cada parto, fração de sangue Gir, fração de sangue Holandês, número de ordenhas por dia (2x,3x) em cada parto e como objetivo de alcance da RNA o parâmetro produção de leite na terceira lactação (PL305\_3). (Figura 3).

A divisão da base de dados para treinamento e validação da rede deu-se em dois conjuntos de dados. A proporção adotada foi de 75% e 25% para treinamento e validação, respectivamente. O primeiro conjunto destina-se ao ajuste de pesos sinápticos e o segundo a fase de avaliação e desempenho da rede (Tabela 4).

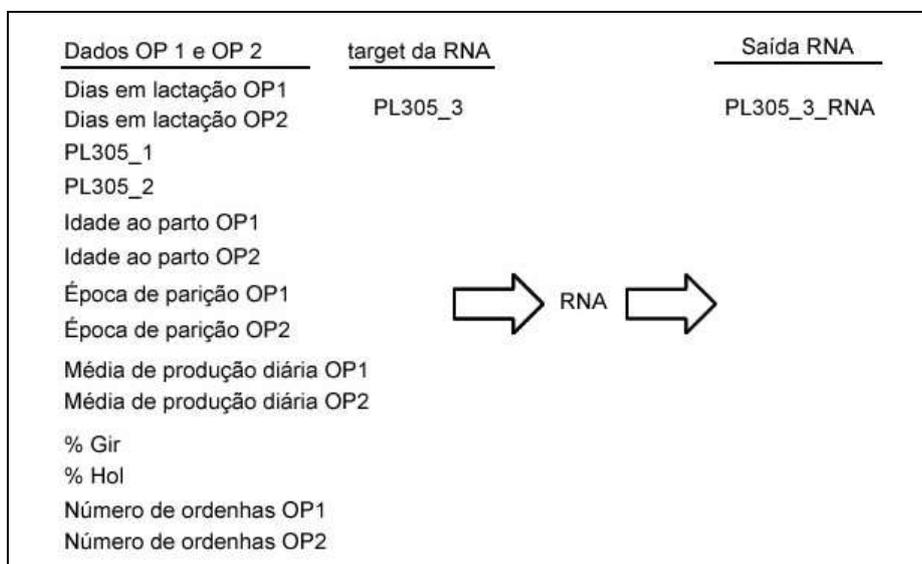


Figura 3. Esquema simplificado da RNA para predição da produção leite no terceiro parto.

Tabela 4. Divisão do conjunto de dados inicial

<i>CATEGORIA</i>	<i>No de dados</i>	<i>%</i>
TREINAMENTO	1227	75
TESTE	409	25

A camada de entrada é composta por quatorze neurônios, sendo estes formados pelos componentes de entrada da rede. A camada oculta é composta por N neurônios. A camada de saída é formada pela estimativa gerada pela RNA para a característica produção de

leite no terceiro parto (PL305\_3\_RNA). (Figura 1)

Para o treinamento da RNA, o número de iterações foi escolhido como critério de parada, de acordo com a Tabela 5, assim como os valores de neurônios ocultos e taxa de aprendizado.

Tabela 5. Propostas de utilização da RNA para predição da produção de leite em 305 dias

<i>Proposta</i>	<i>Camadas ocultas</i>	<i>Neurônios ocultos</i>	<i>Taxa aprendizado</i>	<i>Erro meta</i>	<i>Iterações (Épocas)</i>
A	1	30	0.01	0.0001	1000
B	1	45	0.01	0.0001	1000
C	1	45	0.008	0.0001	1000

### 3. RESULTADOS

#### 3.1 Resultados - Predição da segunda lactação por meio de informações da primeira lactação

Após obter a saída da RNA de acordo com cada proposta de arquitetura elaborada, construiu-se a Tabela 6 que exibe a mensuração do erro quadrático médio de validação para as seis propostas (A - F).

Como se pode observar na Tabela 6, os menores erros quadráticos médios foram alcançados nas propostas D e E (EQM = 0.00001). Entretanto, a proposta E

apresenta número de neurônios pouco maior e valor reduzido da taxa de aprendizado comparada à proposta D. Baixas taxas de aprendizado e elevado número de neurônios acarretam em tempo adicional de processamento dos dados, fato este que justifica a escolha da proposta D como melhor alternativa de predição. Em avaliação visual dos dados, elaborada nas figuras 4 a 10, é nítida a aproximação dos dados de produção reais aos obtidos pela RNA apresentadas nas figuras 7 e 8 (propostas D e E) para dados parciais e figura 10 para dados completos (proposta D).

Tabela 6. Valores do erro quadrático médio (EQM) de acordo com as propostas estabelecidas

<i>Proposta</i>	<i>Camadas ocultas</i>	<i>Neurônios ocultos</i>	<i>Taxa aprendizado</i>	<i>Erro meta</i>	<i>Iterações (Épocas)</i>	<i>EQM Validação</i>
A	1	30	0.001	0.0001	1000	0.00152
B	1	30	0.01	0.0001	1000	0.00011
C	1	35	0.006	0.0001	1000	0.00051
D	1	60	0.01	0.0001	1000	0.00001
E	1	65	0.006	0.0001	1000	0.00001
F	1	65	0.006	0.0001	5000	0.01085

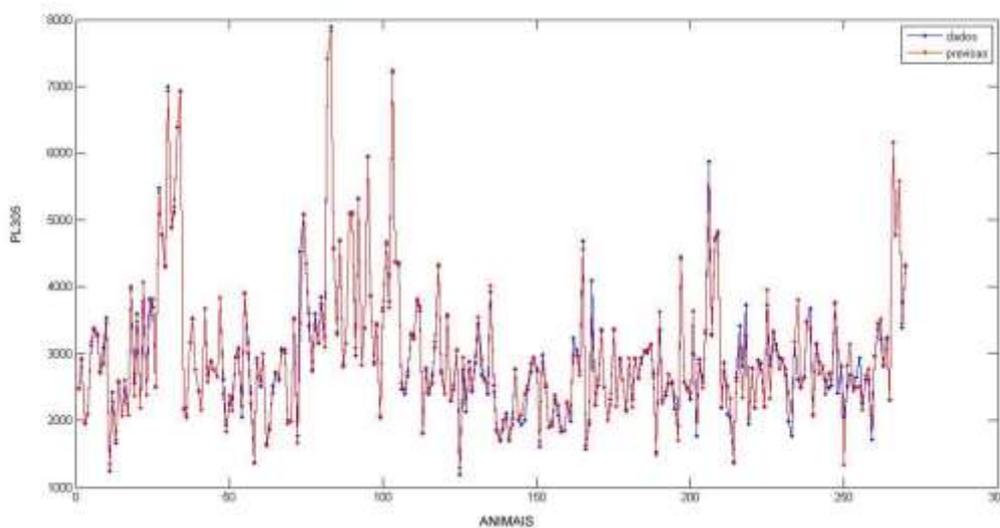


Figura 4. Proposta A (EQM = 0.00152)

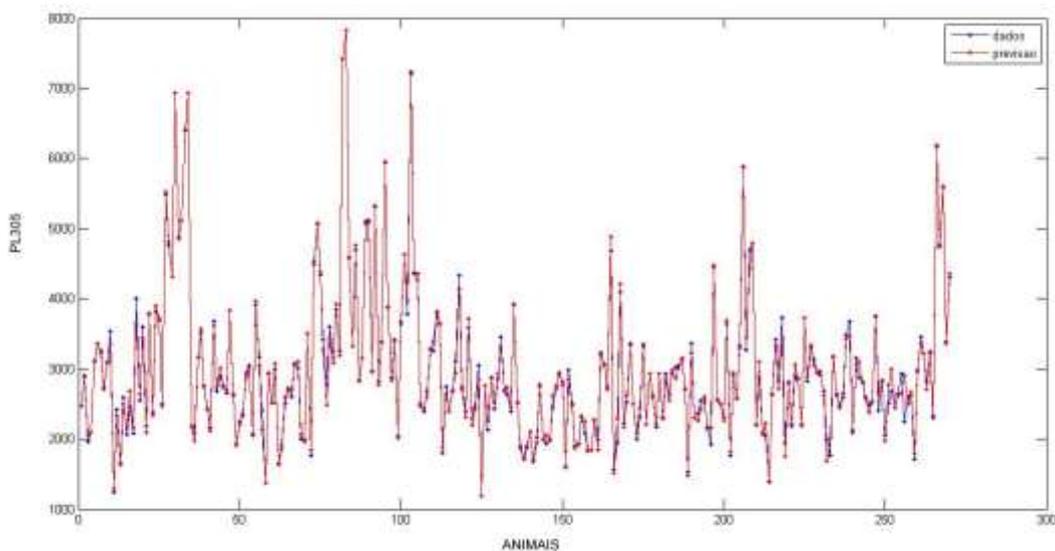


Figura 5. Proposta B (EQM = 0.00011):

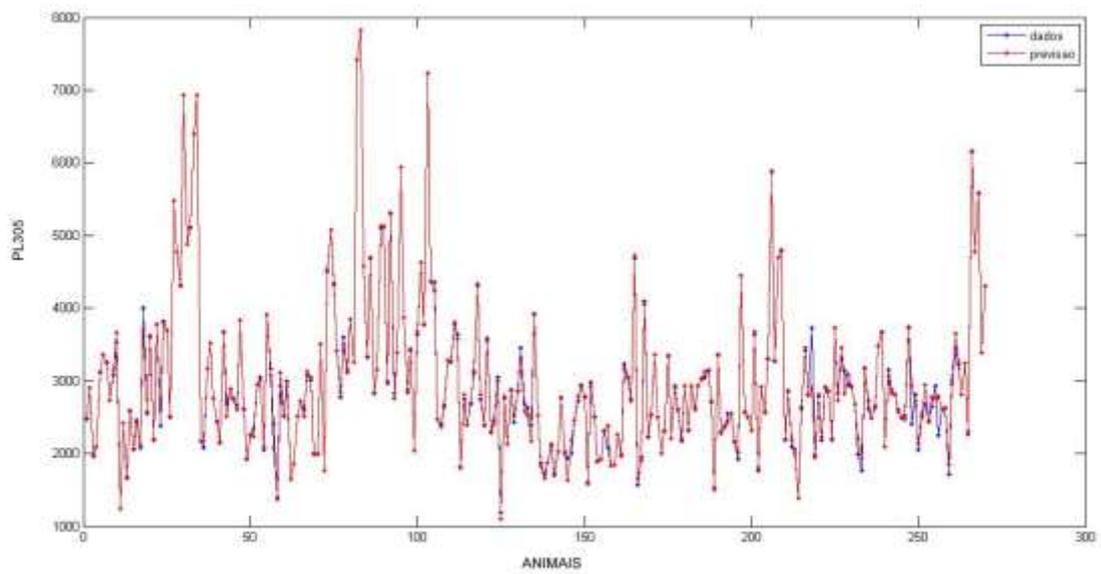


Figura 6. Proposta C (EQM = 0.00051)

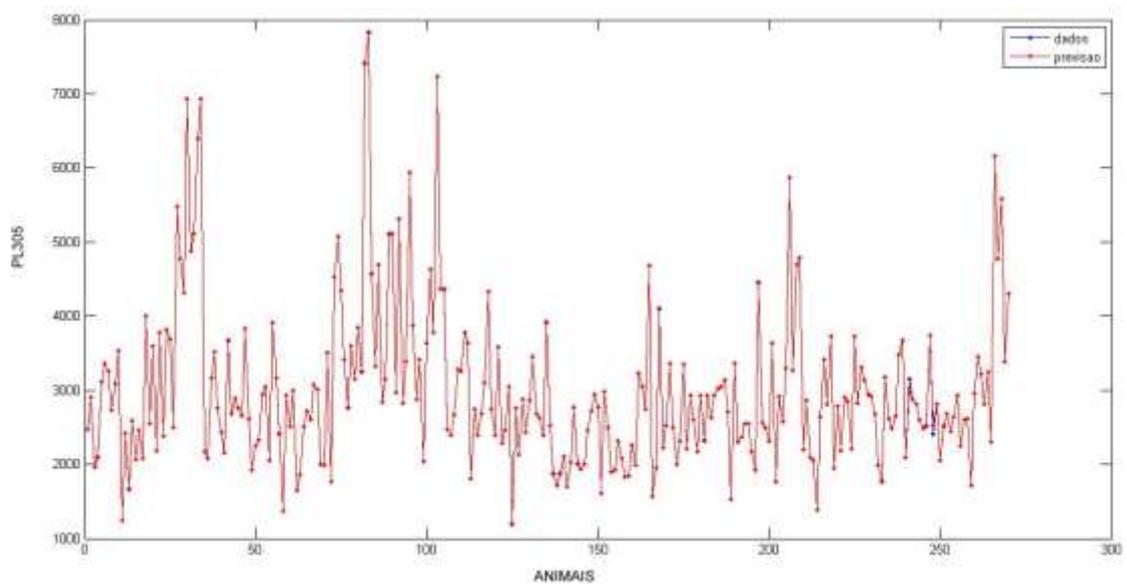


Figura 7. Proposta D (EQM = 0.00001)

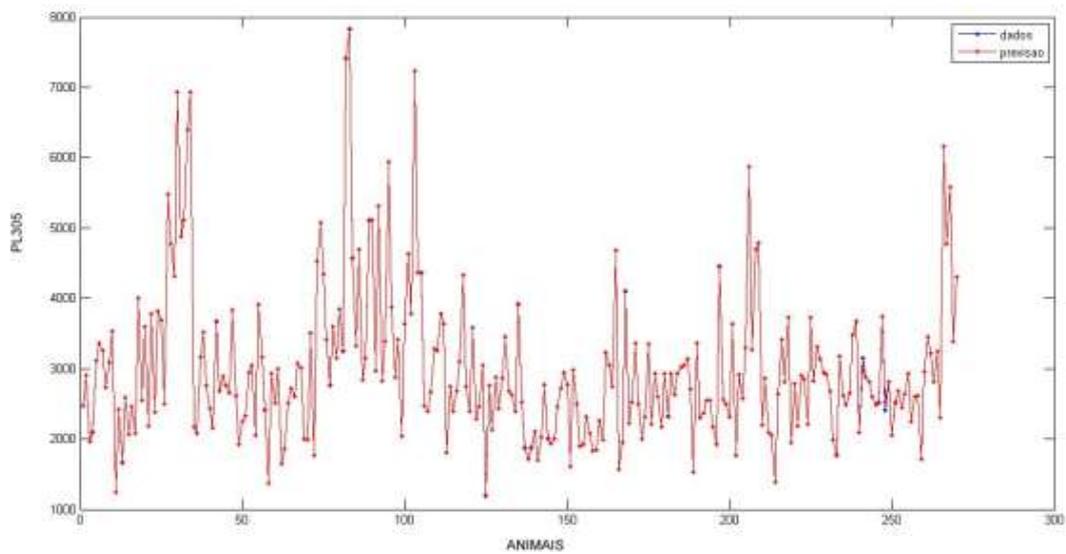


Figura 8. Proposta E (EQM = 0.00001)

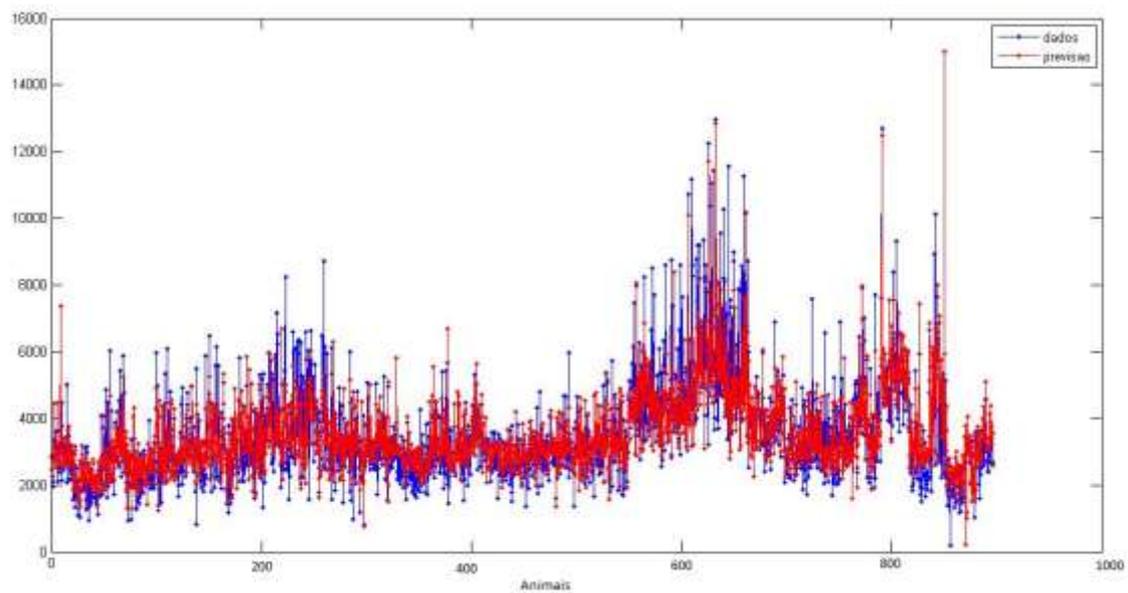


Figura 9. Proposta F para dados integrais sem ajuste ou restrições (EQM = 0.01085)

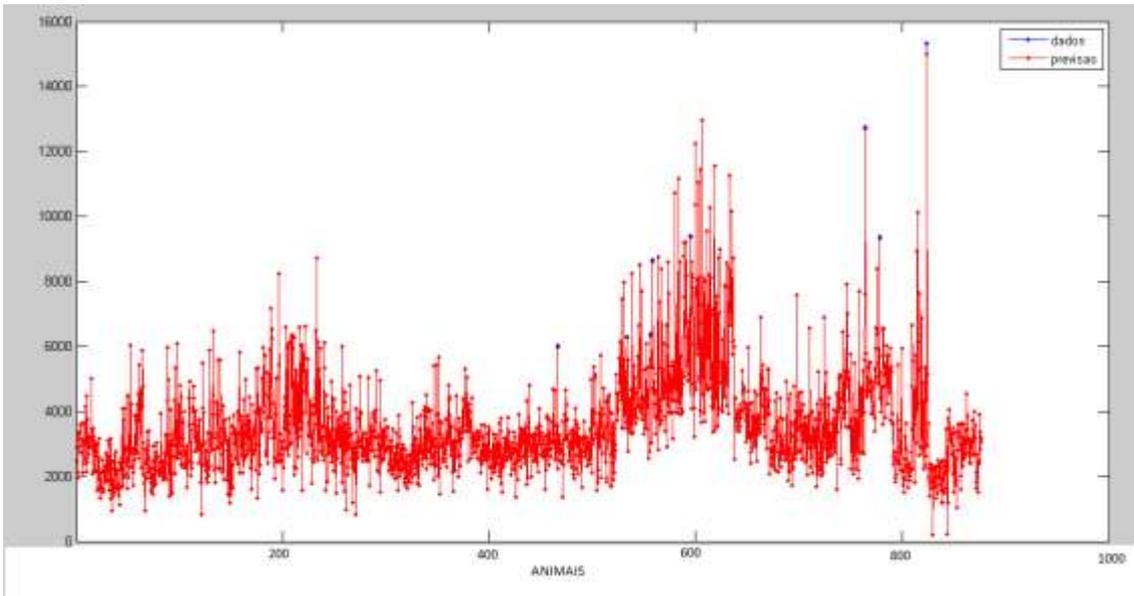


Figura 10. Proposta D para dados completos ajustados (EQM = 0.00001)

Valores da capacidade provável de produção (CPP) foram comparados aos da saída da RNA por meio de correlação de Pearson (r), erro médio de predição e desmembrados de acordo com o grau de participação de genes Gir dos animais avaliados (Tabela 8). A correlação entre a produção real e saída da RNA, nas diferentes propostas de montagem da RNA, é mostrada na Tabela 9.

Da mesma forma a correlação entre a saída RNA e a produção de leite em 305 dias de lactação e o erro médio de predição são apresentados na Tabela 7 e desmembrados de acordo com o grau de participação de genes Gir. A maior correlação ocorreu na proposta D, que também apresentou menor EQM. A proposta F, sem aplicações de restrição no banco de dados apresentou, portanto, valores destoantes da média da população e não obteve treinamento adequado como nas demais propostas que

tiveram a base de dados ajustada para exceções.

O erro médio de predição, obtido pelos quadrados médios das diferenças entre o valor original da produção de leite e saída predita pela RNA, foi obtido da seguinte maneira:

$$\sum (y_i - y_{RNA})^2 / n$$

Em que

$y_i$  = valor original da produção de leite

$y_{RNA}$  = saída RNA

$n$  = número de animais em cada classe

A correlação entre CPP e dados de saída da RNA e a correlação entre CPP e dados de produção de leite em 305 dias de lactação são demonstradas na Tabela 10.

Tabela 7. Correlação entre a saída RNA e produção de leite em 305 dias de lactação de acordo com a participação de genes da raça Gir dos animais avaliados, para os dados de validação.

<i>Fração GIR</i>	<i>Coefficiente Correlação de Pearson(r)</i>	<i>Número de animais (n)</i>	<i>Erro médio de predição</i>
0.125	0,9899	80	17,40
0.25	0,9869	151	11,80
0.3125	0,9901	62	13,90
0.375	0,9733	63	49,81
0.5	0,9968	168	5,01
0.625	0,9897	62	20,35
0.6875	0,9924	36	18,70
0.75	0,9855	106	24,60
0.875	0,9892	37	11,88
1	0,9947	124	6,55

Tabela 8. Correlação entre a CPP e produção de leite em 305 dias de lactação de acordo com a participação de genes da raça Gir dos animais avaliados, para os dados de validação

<i>Fração GIR</i>	<i>Coefficiente Correlação de Pearson(r)</i>	<i>Número de animais (n)</i>	<i>Erro médio de predição</i>
0.125	0,97818	80	20,35
0.25	0,97023	151	14,30
0.3125	0,98198	62	14,70
0.375	0,95908	63	57,81
0.5	0,98237	168	8,67
0.625	0,97025	62	40,15
0.6875	0,98606	36	28,50
0.75	0,97116	106	33,30
0.875	0,9833	37	16,89
1	0,98355	124	7,90

Tabela 9. Correlação de Pearson entre a saída da RNA PL305\_RNA e o valor real da PL305\_2, de acordo com as propostas de montagem da RNA.

<i>PROPOSTA RNA</i>	<i>r (SaídaRNA e PL305_2 original)</i>	<i>Erro médio de predição</i>
A	0.9501	37,45
B	0.9657	28,50
C	0.9689	26,70
D	0.9901	9,15
E	0.9959	6,23
F	0.7900	198,34

Tabela 10. Correlação entre produção de leite em 305 dias e valor mensurado pela CPP e correlação entre a saída da RNA e valor obtido pelo CPP

<i>Fonte de dados</i>	<i>Coefficiente Correlação de Pearson (r)</i>
RNA x CPP	0,9734
PL305 x CPP	0,97150

A predição por meio da capacidade provável de produção (CPP) mostrou-se tão eficaz quanto a RNA, sendo o primeiro método obtido com maior facilidade. A RNA se mostrou eficiente por analisar dados de animais com diferentes participações de genes da raça Gir.

### 3.2 Resultados - Predição da terceira lactação por meio de informações da primeira e segunda lactação

A Tabela 11 foi elaborada após obtenção da saída da RNA, de acordo as propostas de arquitetura demonstradas na Tabela 10 e exibe a mensuração do erro quadrático médio de validação para as três propostas (A - C).

Tabela 11. Estimativas do erro quadrático médio (EQM ou MSE) de acordo com diferentes propostas da RNA

<i>Proposta</i>	<i>Camadas ocultas</i>	<i>Neurônios ocultos</i>	<i>Taxa aprendizado</i>	<i>Erro meta</i>	<i>Iterações (Épocas)</i>	<i>MSE Validação</i>
A	1	30	0.01	0.0001	1000	0.00113
B	1	45	0.01	0.0001	1000	0.00011
C	1	45	0.008	0.0001	1000	0.00001

Como se pode observar na Tabela 11, o menor erro quadrático médio foi alcançado na propostas C (EQM = 0.00001). Para os dados em questão, o menor EQM foi alcançado por meio da menor taxa de aprendizado, justificando seu uso mesmo com tempo de processamento maior do que o obtido na proposta B (em razão da menor taxa de aprendizado). Em avaliação visual dos dados, elaborada nas figuras 11, 12 e 13, observa-se melhor aproximação dos dados de produção reais aos obtidos pela RNA por meio da figura 13 (Proposta C) para dados parciais.

Os dados da Tabela 12 ilustram a semelhança dos dados de saída da RNA em comparação aos dados reais de produção em 305 dias de lactação, para animais com ordem de parto igual ou maior do que três.

A Tabela 13 fornece as correlações de Pearson (r) obtidas entre dados de saída gerados pela RNA e dados reais de produção de leite em 305 dias de lactação. Todas as propostas de RNA apresentaram forte semelhança com resultados dos dados de produção reais, demonstrando a funcionalidade da RNA e o alto reconhecimento de padrão alcançado.

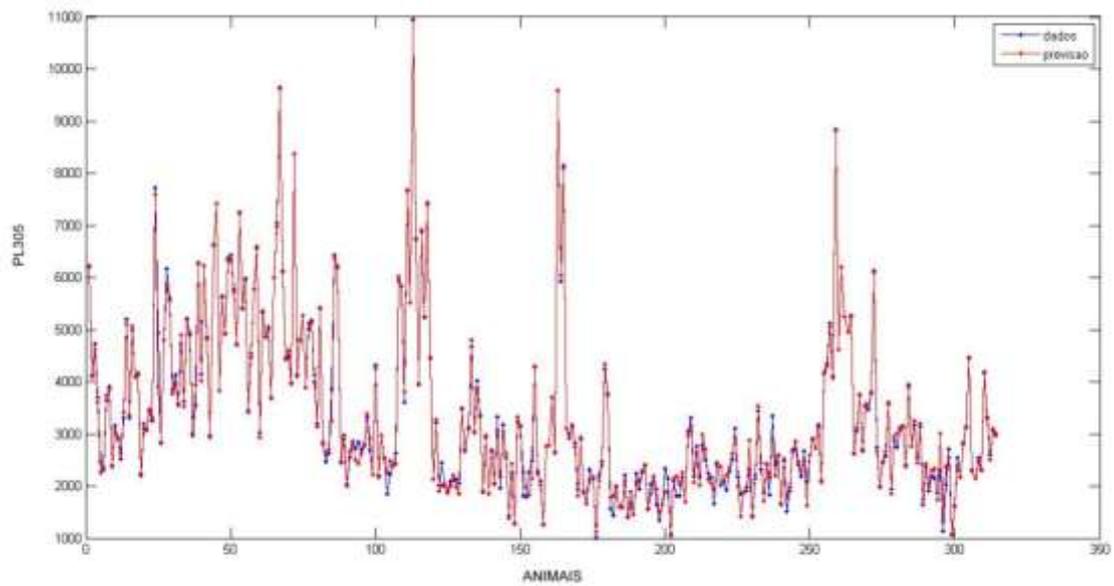


Figura 11. Proposta A após selecionar camada representativa de dados.

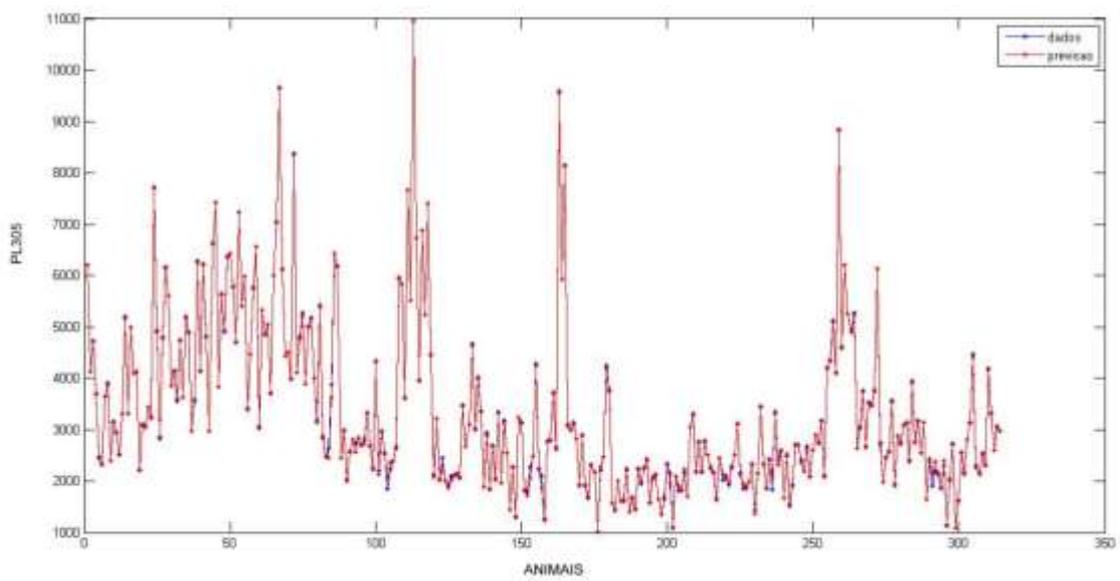


Figura 12. Proposta B após selecionar camada representativa de dados.

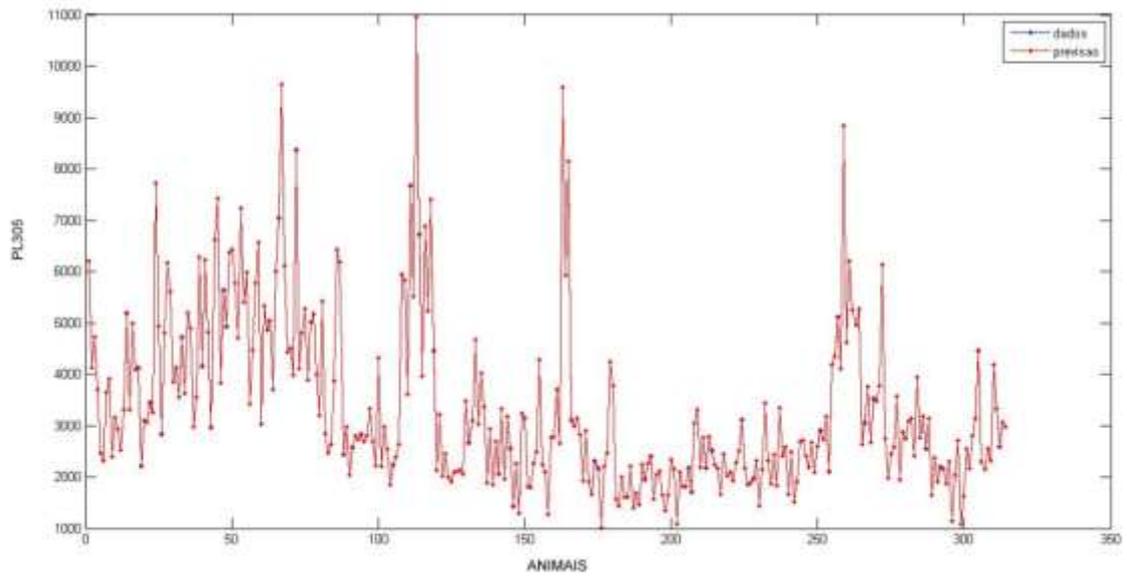


Figura 13. Proposta C após selecionar camada representativa de dados.

Tabela 12. Exemplo de PL305\_3 gerada pela RNA em comparação com PL305\_3 original

<i>Dados ORIGINAIS PL305_3</i>	<i>Dados RNA PL305_RNA</i>
6203	6199
4133	4133
4725	4725
3702	3699
2459	2459
2318	2320
3651	3654
3904	3901
2399	2397
3162	3162
2940	2940
2532	2530
3317	3314
5188	5187
3309	3311
4992	4994
4090	4085
4124	4121
2211	2215
3107	3103
3058	3060
3447	3450
3249	3250
7717	7719
4941	4944

Tabela 13. Correlação de Pearson entre a saída da RNA (PL305\_RNA) e valor real da PL305\_3, de acordo com as propostas de montagem da RNA

<i>PROPOSTA RNA</i>	<i>r (SaídaRNA e PL305_2 original)</i>	<i>Erro médio de predição</i>
A	0.98501	14,60
B	0.98957	13,98
C	0.99590	3,55

#### 4.CONCLUSÕES

As redes neurais artificiais, quando utilizadas em base de dados sem discrepância de valores que caracterizam as mensurações realizadas na trajetória produtiva dos animais, podem obter com precisão mensurações futuras de produção para a característica produção de leite em 305 dias de lactação.

Sistemas baseados em redes neurais artificiais (RNAs) podem ser utilizados como ferramenta para tomada de decisões em propriedades rurais. Esta pode ser utilizada para prever produções futuras em vacas leiteiras de diferentes graus de sangue, baseados no reconhecimento de padrão e relações não lineares descobertos pelas RNAs.

A predição por meio da capacidade provável de produção (CPP) mostrou-se tão eficaz quanto a RNA, sendo a CPP obtida de maneira mais simplificada. A RNA se mostrou eficiente por analisar dados de animais com diferentes participações de genes da raça Gir.

#### 5.REFERENCIAS BIBLIOGRÁFICAS

BERTAZZO, R. P. *Utilização de redes neurais artificiais na predição do mérito*

*genético de bovinos de corte*. Tese (Doutorado) - Faculdade de Zootecnia e Engenharia de Alimentos, Universidade Federal de Minas Gerais, Pirassununga, SP.

BRADLEY, P. S., FAYYAD, U. M.; MANGASARIAN, O. L. Mathematical programming for datamining: formulation and challenges. *NFORMS Journal on Computing*, v.3, n.11, p.217-238, 1998.

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

BRENNECKE, K. *Fracionamento de carboidratos e proteínas e a predição da proteína bruta e suas frações e das fibras em detergentes neutro e ácido de Brachiaria brizantha cv. Mandaru por uma rede neural artificial*. 2007. 138f. Tese (Mestrado em Zootecnia) – Faculdade de Zootecnia e Engenharia de Alimentos, Universidade de São Paulo, Pirassununga, SP.

CAVERO, D.; TOLLE, K. H.; HENZE, C. *et al*. Mastitis detection in dairy cows by application of neural networks.

*Livestock Science*, v.114, n.2-3, p.280-286, 2008.

FALCONER, D. S.; MACKAY, T. F. C. *Introduction to quantitative genetics*. 4.ed. Harlow: Longman, 1996. 464 p.

GRZESIAK, W., P.; BLASZCZYK ; LACROIX, R. Methods of predicting milk yield in dairy cows-predictive capabilities of wood's lactation curve and artificial neural networks (ANNS). *Computers and Electronics in Agriculture*, v.54, n.2, p.69-83, 2006.

HAYKIN, S. *Redes neurais: princípios e prática*. Porto Alegre, RS: Bookman, 2001.

HEALD, C. W.; KIM, T.; SISCHO, W. M. *et al.* A computerized mastitis decision aid using farm – based records: an artificial neural network approach. *Journal of Dairy Science*, v.83, n.4, p.711-720, 2000.

HOSSEINIA, P.; EDRISI, M.; EDRISS, M. A.; NILFOROOSHAN, M. A. Prediction of second parity milk yield and fat percentage of dairy cows based on first parity information using neural network system. *Journal of Applied Science*, v.7, n. 21, p.3274-3279, 2007.

KEPLER FILHO, E. *Melhoramento genético animal no Brasil: fundamentos, história e importância*. Campo Grande: Embrapa, 1999.

LACROIX, R., WADE, K. M.; KOK, R.; HAYES, J. F. Prediction of cow

performance with a connectionist model. *Transactions of the ASAE*, v.38, n.5, p.1573-1579, 1995.

MATLAB guide for personal computers. Version 7.3. Local: MATLAB Institute, 2006.

MEIRELLES, F. D. P. *Modelo computacional de um rebanho bovino de corte virtual utilizando simulação Monte Carlo e redes neurais artificiais*. 2005. 104f. Tese (Doutorado em Zootecnia) – Faculdade de Zootecnia e Engenharia de Alimentos, Universidade de São Paulo, Pirassununga, SP.

MENDES, R. A.; BERTAZZO, R. P.; SILVA, D. C. et al. Sistemas especialistas na predição do fenótipo peso ao sobreano em bovinos da raça Brahman. In: SIMPÓSIO INTERNACIONAL DE INICIAÇÃO CIENTÍFICA DA USP (SIICUSP), 2009, Pirassununga SP. *Anais...* Pirassununga,: USP, 2009.

MEYER, K. WOMBAT: digging deep for quantitative genetic analyses by restricted maximum likelihood. In: WORLD CONGRESS ON GENETIC APPLIED TO LIVESTOCK PRODUCTION, 8., 2006, Belo Horizonte. *Proceedings...* Belo Horizonte: Instituto Prociência, 2006. CD-ROM.

NAZARIO, S. L. S.; KITANO, C.; HIGUTI, R. T. et al. Use of ultrasound neural networks to detect adulteration in milk samples.. In: INTERNACIONAL

CONGRESS ON ULTRASONICS, 2007, Viena, Itália. *Anais...* Viena: Viena University of Technology, 2007. p.9-12.

NEVES, A.R.M. *Aplicação de redes neurais artificiais na predição de valores genéticos em bovinos de leite da Raça Pardo-Suiça*. Belém: Universidade Federal do Pará, Instituto de Tecnologia, 2007.

NJUBI, D. M.; WAKHUNGU, J. W.; BADAMANA, M. S. Use of test-day records to predict first lactation 305-day milk yield using artificial neural network in Kenyan Holstein–Friesian dairy cows. *Tropical Animal Health and Production*, v.42, n.4, p.639-644, 2009.

PEREIRA, J. C. *Melhoramento genético aplicado à produção animal*. Belo Horizonte, FEP-MVZ, 2004. 610p.

PEREIRA, R. J. *Medidas de persistência da produção de leite em rebanhos Gir Leiteiro utilizando modelos de regressão aleatória*. Viçosa: Universidade Federal de Viçosa, 2009.

SABLANI, S. S.; RAMASWARNY, H. S.; PRASHER, S. O. A neural network approach for thermal processing

applications. *Journal of Food Processing and Preservation*, v.19, n.4, p.283-301, 1995.

SALEHI, F., LACROIX, R.; WADE, K. M. Effects of learning parameters and data presentation on the performance of back-propagation networks for milk yield prediction. *Transactions of the ASAE*, v.41, n.1, p.253-259, 1998.

SAMARASINGHE, S. *Neural networks for applied sciences and engineering: from fundamentals to complex pattern recognition*. Boca Raton, FL: CRC Press, 2006.

SUN, Z.; SAMARASINGHE, S.; JAGO, J. Detection of mastitis and its stage of progression by automatic milking systems using artificial neural networks. *Journal of Dairy Research*, v.77, n.2, p.168-175, 2010.

YANG, C. C.; PRASHER, S. O.; LANDRY, J. A. *et al.* Application of artificial neural networks in image recognition and classification of crop and weeds. *Canadian Agricultural Engineering*, v.42, n.3, p.147-152, 2000.