



FLÁVIA CRISTINA MARTINS QUEIROZ MARIANO

**REDES NEURAIIS NA PREDIÇÃO DE
VALORES ENERGÉTICOS DE
ALIMENTOS PARA AVES**

**LAVRAS - MG
2014**

FLÁVIA CRISTINA MARTINS QUEIROZ MARIANO

**REDES NEURAIS NA PREDIÇÃO DE VALORES ENERGÉTICOS DE
ALIMENTOS PARA AVES**

Tese apresentada à Universidade Federal de Lavras, como parte das exigências do Programa de Pós-Graduação em Estatística e Experimentação Agropecuária, Área de concentração em Estatística e Experimentação Agropecuária, para a obtenção do título de Doutor.

Orientador
Prof. Dr. Renato Ribeiro de Lima

**LAVRAS - MG
2014**

**Ficha Catalográfica Elaborada pela Coordenadoria de Produtos e
Serviços da Biblioteca Universitária da UFLA**

Mariano, Flávia Cristina Martins Queiroz.

Redes neurais na predição de valores energéticos de alimentos
para aves / Flávia Cristina Martins Queiroz Mariano. – Lavras :
UFLA, 2014.

128 p. : il.

Tese (doutorado) – Universidade Federal de Lavras, 2014.

Orientador: Renato Ribeiro de Lima.

Bibliografia.

1. Comitê de redes neurais. 2. Energia metabolizável. 3.
Intervalos HPD. 4. Predição. 5. Porcentagem de acerto. I.
Universidade Federal de Lavras. II. Título.

CDD – 519.536

FLÁVIA CRISTINA MARTINS QUEIROZ MARIANO

**REDES NEURAIS NA PREDIÇÃO DE VALORES ENERGÉTICOS DE
ALIMENTOS PARA AVES**

Tese apresentada à Universidade Federal de Lavras, como parte das exigências do Programa de Pós-Graduação em Estatística e Experimentação Agropecuária, Área de concentração em Estatística e Experimentação Agropecuária, para a obtenção do título de Doutor.

APROVADA em 28 de janeiro de 2014.

Profa. Dra. Thelma Sáfadi	DEX-UFLA
Prof. Dr. Antônio Policarpo Souza Carneiro	DET-UFV
Prof. Dr. Joel Augusto Muniz	DEX-UFLA
Prof. Dr. Paulo Borges Rodrigues	DZO-UFLA
Prof. Dr. Wiliam Soares Lacerda	DCC-UFLA



Prof. Dr. Renato Ribeiro de Lima
Orientador

**LAVRAS - MG
2014**

*Ao meu marido, Jônatas, pelo amor sincero e por ser meu companheiro,
incentivando-me a percorrer este caminho, compartilhando angústias e
dúvidas em momentos difíceis.*

*Aos meus pais, Ovídio e Ana Cristina Queiroz, pelo amor incondicional,
conselhos, incentivo e apoio em cada instante da minha vida;*

*Aos meus irmãos, Flaviano e Frederico. Aos meus sobrinhos, cunhadas, sogros,
demais familiares e amigos. A todos estes que, mesmo estando a alguns
quilômetros de distância, sempre me cercaram de carinho, zelo e motivação.*

DEDICO E OFEREÇO

AGRADECIMENTOS

A Deus, o autor da minha fé, pela sua presença constante, permitindo-me e capacitando-me a concluir mais esta etapa em minha vida.

Ao meu marido, Jônatas Marcon Mariano, pelo seu amor e paciência em todos os momentos. A conclusão desta etapa é apenas mais uma de nossas conquistas conjuntas. Obrigada por sempre acreditar em mim... te amo!!!!

Aos meus pais, imprescindíveis em minha vida, por todo amor e apoio a mim dedicados. Vocês são meus exemplos de vida.

Aos meus irmãos, Flaviano e Frederico, pelo eterno amor fraternal. Aos meus sobrinhos, sogros, cunhadas e cunhados que me proporcionam momentos de felicidade e descontração.

À minha super amiga Stefani, por me ouvir e me aconselhar sempre.

Ao meu orientador, o professor Dr. Renato Ribeiro de Lima, pelos valiosos ensinamentos tanto acadêmicos quanto pessoais, pela orientação, compreensão e confiança durante estes anos de parceria.

À Universidade Federal de Lavras e ao Departamento de Ciências Exatas, pela oportunidade de cursar o doutorado e por acreditarem em meu potencial.

Aos professores do Programa de Pós-Graduação em Estatística e Experimentação Agropecuária, do Departamento de Ciências Exatas, pelos ensinamentos. Em especial, aos professores Dr. Daniel Ferreira Furtado, Dr. Júlio Sílvio de Sousa Bueno Filho e aos colegas, Dr. Walmes Marques Zeviani e Dr. Diogo Francisco Rossoni, pelo auxílio em eventuais dúvidas.

Aos companheiros, Dr. Crystian Arantes Paixão e Dra. Renata Ribeiro Alvarenga, pela parceria e pelos conhecimentos compartilhados.

A todos os funcionários do Departamento de Ciências Exatas, pela atenção e

pelo bom atendimento sempre que necessário.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), pela concessão da bolsa de estudos, durante o período de realização deste trabalho.

Às parceiras de apartamento, Marília e Jessica, pela companhia e amizade.

Às minhas amigas de estudo - Deyse, Franciella e Rosi - pela amizade e, também, aos demais colegas de curso, por compartilhar conhecimentos e experiências acadêmicas.

Aos membros da banca do exame de qualificação e da banca examinadora da defesa de tese. A estes agradeço pela disponibilidade, pelos valiosos comentários e sugestões para melhorar a qualidade deste trabalho.

Enfim, a todos aqueles familiares e amigos que, de alguma maneira, contribuíram para que este percurso pudesse ser concluído.

Obrigada a todos.... Deus os abençoe!!!

“Confia no Senhor de todo o teu
coração, e não te estribes no teu
próprio entendimento.”

Provérbios 3:5

“Porque Dele, e por Ele, e para Ele,
são todas as coisas; glória, pois, a
Ele eternamente. Amém.”

Romanos 11:36

RESUMO GERAL

Na formulação de rações equilibradas nutricionalmente é necessário um conhecimento preciso da composição química e dos valores energéticos dos alimentos. Este conteúdo energético pode ser determinado utilizando bioensaios, tabelas de composição nutricional, equações de predição baseadas na composição química dos alimentos e, recentemente, têm sido implementados modelos de redes neurais artificiais (ANN). Contudo, os trabalhos envolvendo ANN não consideraram incerteza nas predições, nem trabalharam com comitês de redes. Neste trabalho, inicialmente, foi proposta a implementação de um modelo de ANN para predição da energia metabolizável aparente corrigida pelo balanço de nitrogênio (EMAn) de alimentos para aves de corte. Com base nos resultados obtidos neste modelo de redes neurais, foi criada uma ferramenta de cálculo, permitindo que os nutricionistas de animais obtenham uma rápida e acurada predição de EMAn. Um comitê de redes neurais *perceptron* multicamadas (*MLP*), também, foi implementado para prever a EMAn, com o objetivo principal de considerar incertezas nas predições deste comitê, considerando intervalos de credibilidade de máxima densidade de probabilidade (HPD) para os alimentos do conjunto de teste. Para comparar, em termos quantitativos, o modelo de comitê de redes com modelo de regressão linear múltipla, foi definida a porcentagem de acerto de predição, considerando os intervalos de predições determinados em cada modelo. Além disso, a acurácia de cada modelo selecionado foi avaliada com base em medidas de erro entre os valores observados e preditos. Dentre os modelos selecionados, o comitê de redes, em que 1000 redes foram combinadas pelo estimador da moda da distribuição empírica das predições, garantiu uma maior acurácia nas predições. Portanto, este comitê será disponibilizado para a predição da EMAn de alimentos concentrados energéticos e proteicos comumente utilizados na formulação de rações avícolas.

Palavras-chave: Comitê de Redes Neurais. Energia Metabolizável. Intervalos HPD. Predição. Porcentagem de Acerto.

GENERAL ABSTRACT

The precise knowledge of chemical composition and food energy values is necessary to formulate nutritionally balanced rations. The energetic content may be determined using bioassays, nutritional composition tables and prediction equations based on food chemical composition. Recently, artificial neural network (ANN) models have been implemented to predict metabolizable energy of poultry feedstuffs. However, neither a committee of networks nor the uncertainties related to the predictions of the metabolizable energy were considered. In this work, an ANN model was implemented to predict the nitrogen-corrected apparent metabolizable energy values (AMEn) of feedstuffs for broilers. One Excel macro was created based on the results of the best ANN model, which allows an efficient way to predict the AMEn values. Another objective of this work was to quantify the uncertainty related to the predicted values of AMEn, considering Highest Probability Density (HPD) credibility intervals obtained by using a committee of neural networks. A test dataset, which was obtained from bioassays, was used to evaluate the percentage of correct predictions (PCP) of different feedstuffs. The PCP was used as quantitative criterion to compare the predictions obtained using the network committee and the multiple linear regression models. Furthermore, the accuracy of the models was evaluated based on their values of error measures between the observed and predicted values. The results of this study demonstrated that, the accuracy of the predicted values of AMEn by considering the mode of an empirical distribution, obtained from 1000 networks of a committee, was higher than that obtained by using the regression model. Thus, the network committee is recommended as the most accurate model to predict the AMEn values of the feedstuffs energetic concentrates used by the poultry feed industry.

Keywords: HPD Intervals. Metabolizable Energy. Neural Networks Committee. Prediction. Percentage of success.

LISTA DE ABREVIATURAS

ADF	acid detergent fiber
AMEn	apparent metabolizable energy corrected for nitrogen balance
ANN	redes neurais artificiais
Bias	viés (erro médio)
CAT	categoria do alimento
CF	crude fiber
CN	committee of networks
CP	crude protein
EE	extrato etéreo
EM	energia metabolizável
EMAn	energia metabolizável corrigida pelo balanço de nitrogênio
FB	fibra bruta
HPD	máxima densidade de probabilidade
ING	tipo de ingrediente
LM	Levenberg-Marquardt
m.s.e.	erro quadrático médio
MAD	desvio médio absoluto
MAPE	erro médio percentual absoluto
ME	metabolizable energy
MLP	<i>perceptron</i> multicamadas

MM matéria mineral
NDF neutral detergent fiber
PB proteína bruta
PCP porcentagem de acerto da predição
 R^2 coeficiente de determinação
RLMP regressão linear múltipla ponderada
RPROP Resilient backpropagation

SUMÁRIO

	PRIMEIRA PARTE	18
1	INTRODUÇÃO	18
1.1	Estrutura da tese	20
2	REFERENCIAL TEÓRICO	21
2.1	Modelos de predição de valores energéticos de alimentos para aves	21
2.2	Meta-análise	24
2.3	Redes neurais artificiais	26
2.4	Comitê de redes neurais	47
2.5	Critérios de adequabilidade de ajuste.	51
2.6	Porcentagem de acerto da predição - PCP	53
2.7	Intervalos de máxima densidade de probabilidade - HPD	54
3	SOFTWARES UTILIZADOS.	57
	REFERÊNCIAS.	59
	SEGUNDA PARTE - ARTIGOS.	65
	ARTIGO 1 Prediction of the energy values of feedstuffs for broilers using meta-analysis and neural networks.	66
	ARTIGO 2 Neural network committee to predict the AMEn of poultry feedstuffs	85
	ARTIGO 3 Comitê de redes neurais artificiais e regressão múltipla ponderada para a predição de valores energéticos de alimentos para aves de corte.	109
	CONSIDERAÇÕES GERAIS	127

PRIMEIRA PARTE

1 INTRODUÇÃO

Em razão da diversidade de alimentos e seus subprodutos, utilizados na formulação de rações para aves, surge a necessidade de se conhecerem bem seus valores nutritivos e energéticos, para melhor aproveitamento e utilização dos mesmos. Por causa dessa diversidade, o sucesso na formulação de rações requer um conhecimento preciso da composição química e dos valores de energia metabolizável (EM) dos alimentos, o que possibilita um fornecimento adequado de energia para os animais. Porém, a determinação da energia metabolizável não é de fácil execução, além de demandar tempo e ser de custo elevado.

Uma forma rápida e menos onerosa de se determinar esses valores é a utilização de equações de predição, estabelecidas em função da composição química dos mesmos. Existem à disposição na literatura vários trabalhos de pesquisa, onde se estabeleceram equações de predição para os valores energéticos de uma série de alimentos. Em alguns destes trabalhos (BORGES et al., 2003; RODRIGUES et al., 2001, 2002; ZHAO et al., 2008) foi utilizada a análise de regressão múltipla, sem qualquer tipo de restrição ou ponderação, determinando resultados inconsistentes (ALVARENGA et al., 2011).

Visando obter resultados mais consistentes, em alguns estudos foram determinadas equações de predição para a energia metabolizável aparente corrigida pelo balanço de nitrogênio (EMAn) de alimentos para aves, utilizando meta-análise. Nesses estudos, os grupos homogêneos de resultados experimentais foram obtidos pela combinação de níveis de fatores (NASCIMENTO et al., 2009, 2011) ou pelo uso da técnica multivariada de componentes principais (MARIANO et al., 2012). Nestas equações apresentaram-se resultados semelhantes aos disponíveis na literatura sem meta-análise, porém com um baixo custo em decorrência da utilização da

meta-análise em vez de executarem experimentos específicos como em Rodrigues et al. (2001, 2002). Além disso, o desenvolvimento da meta-análise foi facilitado em Mariano et al. (2012), uma vez que não houve a necessidade de definir fatores mais adequados para a formação de grupos homogêneos de resultados experimentais, como em Nascimento et al. (2009, 2011).

Outra metodologia que, recentemente, tem sido utilizada como uma forma de predição para EM são as redes neurais artificiais (ANN). Porém, nos estudos encontrados (AHMADI et al., 2008; AHMADI; MOTTAGHITALAB; NARIMANZADEH, 2007; PERAI et al., 2010), além de serem utilizados poucos dados na predição de energia metabolizável de alimentos para aves, foram consideradas poucas variáveis de entrada e no máximo dois tipos de alimentos. Assim, as redes neurais propostas foram específicas para determinados tipos de alimentos.

Realizou-se este trabalho com os objetivos:

- considerar na meta-análise um método computacional de ANN para estimar os valores de EMAn de diferentes alimentos como milho, farelo de soja, e subprodutos destes, dentre outros comumente utilizados na formulação de rações para aves de corte;
- criar uma ferramenta de cálculo baseada em resultados de redes neurais, permitindo que os nutricionistas da área animal obtenham uma rápida e acurada predição de EMAn;
- considerar incerteza nas predições de modelos de ANN, com o uso de um comitê de redes para a predição da EMAn;
- comparar o modelo de comitê de redes com equação de predição obtida por metodologia estatística anteriormente recomendada na literatura.

Para avaliar esta nova proposta, foi utilizado o maior número possível de

variáveis de entrada que interferem na variabilidade da EMAn, o que nem sempre é viável em um modelo de regressão múltipla. Dessa forma, buscou-se obter um modelo mais acurado do que os obtidos em equações de predição, utilizando ou não o princípio da meta-análise.

1.1 Estrutura da tese

Inicialmente será apresentada uma revisão bibliográfica sobre modelos de predição de alimentos para aves, meta-análise, redes neurais e comitê de redes. Na abordagem sobre redes, será dada uma ênfase especial à rede neural *perceptron* multicamadas (*MLP*) e aos algoritmos de treinamento *backpropagation* e de Levenberg-Marquardt, por serem eles utilizados neste trabalho. Também é apresentada a medida estabelecida para a avaliação da taxa de acerto de predição e alguns critérios de adequabilidade de modelos. Na segunda parte, estão anexados três artigos relacionados ao desenvolvimento desta tese, os quais estão redigidos, conforme normas exigidas por cada um dos periódicos científicos que foram selecionados para publicação. Ao final, serão apresentadas as considerações gerais, com base nos resultados obtidos durante o desenvolvimento deste trabalho de tese.

No primeiro artigo, estão descritos os resultados da aplicação da rede *MLP* na predição da EMAn de alimentos para aves. Este artigo foi publicado no periódico *Animal* (Cambridge). No segundo artigo, foi proposto um comitê de redes neurais, visando considerar incertezas nas predições da EMAn por meio de intervalos de credibilidade HPD. Este segundo artigo foi submetido ao periódico *NC&A* (Neural Computing and Applications). No terceiro artigo, foi definida uma medida de porcentagem de acerto de predição, com base em intervalos de predição, para a comparação de modelos de comitê de redes e de equações de predições. Este último artigo será submetido à revista *Ciência Rural*.

2 REFERENCIAL TEÓRICO

2.1 Modelos de predição de valores energéticos de alimentos para aves

O sucesso na formulação de rações de mínimo custo requer um conhecimento preciso da composição química e dos valores de energia metabolizável dos alimentos, o que possibilita um fornecimento adequado de energia para os animais (ALBINO et al., 1992; NASCIMENTO et al., 2011). Além de ser indispensável para as funções vitais das células, a energia tem participação importante na regulação do consumo, sendo ela que permite que o animal ganhe peso, desenvolva-se bem, renda lucros, etc.

A determinação do valor da EM pode ser obtida por meio de ensaios biológicos (WAN et al., 2009; ZHAO et al., 2008), o que não é de fácil execução, além de demandar tempo e ser de custo elevado. Uma outra forma, para obtenção desses valores, é o uso de tabelas de composição de alimentos. Contudo, torna-se inseguro para a indústria utilizar os valores de tabela, em virtude da diversidade de valores que as mesmas sugerem por causa da diversidade de variedades, processamento e armazenamento do grão, condições de plantio, regiões geográficas, entre outros (NAGATA et al., 2004; NASCIMENTO et al., 2011).

No entanto, existe uma forma rápida de se determinar os valores energéticos dos alimentos, que consiste na utilização de equações de predição, estabelecidas em função da composição química dos alimentos, normalmente de fácil e rápida obtenção (ALBINO et al., 1992; NASCIMENTO et al., 2011).

A vantagem de se determinar equações de predição da energia metabolizável, em função da composição química dos alimentos, decorre de que é bem mais fácil realizarem-se análises laboratoriais para determinação das composições químicas dos alimentos, do que efetuar bioensaios com os animais e ficar medindo

suas excretas em calorímetros. Além disso, a maioria dos laboratórios, indústrias e estações de pesquisa do país nem sempre possuem calorímetros e, mesmo que possuíssem, a execução desses ensaios biológicos é dispendiosa e demorada. Nesse caso, trabalhar com equações geradas com base em análises de composições químicas do alimento pode auxiliar o zootecnista (ALBINO et al., 1992; MARIANO et al., 2012; NASCIMENTO et al., 2011).

Existem, na literatura nacional, alguns trabalhos estabelecendo equações para a predição de valores energéticos de uma série de alimentos. Utilizando os valores de EMAn, determinados em ensaios metabólicos, pelo método tradicional de coleta total de excretas com pintos em crescimento, foram obtidas e aplicadas em validações, equações para alimentos energéticos (NAGATA et al., 2004; RODRIGUES et al., 2001) e para alimentos proteicos (RODRIGUES et al., 2002; ZONTA et al., 2004), com o objetivo de se predizer a energia como função da composição química dos alimentos. Nestes trabalhos, utilizou-se uma análise de regressão múltipla, sem qualquer tipo de restrição ou ponderação. Porém, os resultados obtidos, utilizando-se essas equações, não têm sido satisfatórios ou são aplicadas a um só tipo de alimento.

Para obter resultados mais consistentes, uma maneira interessante é trabalhar com a meta-análise, ou seja, utilizando procedimentos estatísticos, combinam-se vários resultados de estudos distintos, mas relacionados entre si (SAUVANT et al., 2008). Na meta-análise, embora os estudos envolvam um mesmo assunto, é importante que grupos mais homogêneos possíveis sejam formados, visando controlar a heterogeneidade existente entre eles (LOVATTO et al., 2007; MARIANO et al., 2012).

A aplicação da meta-análise, para a obtenção de equações de predição da energia metabolizável de alimentos, pode ser feita de diferentes maneiras. Em

alguns estudos disponíveis na literatura, como em Nascimento et al. (2009, 2011), os grupos homogêneos de resultados experimentais, utilizados em equações de predição de EMAn, foram obtidos por meio da combinação dos níveis de três fatores que influenciavam na variabilidade da EMAn (sexo, idade e metodologia empregada) ou como em Mariano et al. (2012), com a técnica multivariada de componentes principais.

Modelos de redes neurais artificiais (ANN), também, têm sido utilizados para prever valores energéticos de alimentos (AHMADI et al., 2008; AHMADI; MOTTAGHITALAB; NARIMAN-ZADEH, 2007; MARIANO et al., 2013; PERAI et al., 2010). Porém, em Ahmadi et al. (2008), Ahmadi, Mottaghitlab e Nariman-Zadeh (2007) e Perai et al. (2010) foram considerados poucos dados e poucas variáveis de entrada no desenvolvimento da rede neural para predição da EM de, no máximo, dois tipos de alimentos para aves. Além disso, a rede foi desenvolvida para alimentos específicos. Em Mariano et al. (2013), a rede *MLP* foi aplicada com sucesso para prever a EMAn de diferentes alimentos (milho, farelo de soja e subprodutos destes, dentre outros) comumente utilizados na formulação de aves de corte, considerando o maior número possível de variáveis de entrada que interferem na variabilidade da EMAn. Contudo, o modelo de ANN obtido por Mariano et al. (2013) é, ainda, mais apropriado para estimar a EMAn de alimentos tais como o milho, sorgo, soja, farelo de soja, trigo e farinha de trigo, em razão da maior representatividade que estes alimentos tiveram no conjunto de dados utilizados no desenvolvimento do modelo.

A vantagem de se prever valores energéticos pelo uso de ANN é que, nas redes, é possível utilizar um maior número de variáveis que venham interferir na variabilidade da EM, o que nem sempre é viável em um modelo de regressão múltipla. Além do que, a aplicação prática de modelos de regressão mais complexos

torna-se inviável para os nutricionistas da área animal. Assim, com o uso de modelos de ANN, espera-se obter resultados mais acurados em razão do maior controle da variabilidade. Portanto, uma outra forma de prever a EMAn é considerar uma meta-análise utilizando modelos de redes neurais artificiais.

No entanto, a comparação de diferentes modelos de predição de valores energéticos, assim como validá-las posteriormente, é de suma importância. Dessa forma, além do desenvolvimento teórico, com a obtenção de modelos de predição de uma forma mais acurada, é muito importante que se faça a validação dos resultados obtidos pelos ensaios metabólico e de desempenho. Esta validação indicará se os métodos utilizados são, realmente, melhores do que os utilizados até então (ALVARENGA et al., 2011; MARIANO et al., 2012).

2.2 Meta-análise

A repetição ou confirmação de resultados experimentais tem grande importância, pois espera-se que um resultado obtido, em determinado trabalho científico, confirme-se com a realização de outros trabalhos semelhantes. Porém, essa premissa nem sempre é verdadeira e, em muitos casos, obtêm-se variações nos resultados que colocam em dúvida as conclusões obtidas.

Uma maneira interessante de se trabalhar com os vários resultados obtidos seria utilizar uma metodologia que permitisse a combinação desses diferentes resultados e, assim, obter conclusões que, inicialmente, parecem ser mais confiáveis e informativas.

A meta-análise é definida como o uso de métodos estatísticos para combinar ou comparar resultados de estudos distintos, mas relacionados (SAUVANT et al., 2008), isto é, a meta-análise é a “análise das análises”. Esses métodos estatísticos empregados na meta-análise asseguram a obtenção de uma estimativa

combinada precisa e acurada, sobretudo em virtude do aumento do número de observações e, conseqüentemente, do poder estatístico e da possibilidade de examinar a variabilidade entre os estudos (FAGARD; STAESSEN; THIJS, 1996).

Um dos primeiros trabalhos cujo autor se interessou pelo problema de combinar resultados de vários experimentos independentes é o de Cochran (1954). Eysenck, em 1952, concluiu, em um estudo, que a psicoterapia que nada mais é que o tratamento de algum distúrbio psíquico, não tinha nenhum efeito benéfico, surgindo, com isso, uma polêmica entre os profissionais da área. Mas só depois de 20 anos de pesquisa e centenas de estudos, tentando resolver esse dilema que, Smith & Glass em 1977, considerando 375 pesquisas em psicoterapia, conseguiram concluir que psicoterapia tem efeitos sim e provaram, então, que Eysenck estava errado (LOVATTO et al., 2007). O método utilizado por eles, para combinar resultados de diferentes estudos, era conhecido como meta-análise.

O principal objetivo da meta-análise é a obtenção de novos resultados, considerando a síntese de observações distintas, porém relacionadas entre si. Assim, avalia-se um número maior de trabalhos, o que garante a obtenção de um resultado mais consistente e acurado. Conseqüentemente, o poder analítico é melhorado, aumentando as chances de identificar diferenças entre os tratamentos, se as mesmas existirem. Além disso, a meta-análise tem por objetivo sintetizar resultados contraditórios, possibilitando a generalização de conclusões para uma amplitude variada de estudos, assim como resolver incertezas, quando certas conclusões se destoam (LOVATTO et al., 2007; MARIANO et al., 2012).

As etapas necessárias ao desenvolvimento de uma meta-análise são: a) definição do objetivo do trabalho, do tema da pesquisa; b) sistematização das informações; c) codificação dos dados; d) filtragem dos dados; e) análise dos dados incluídos na base; f) formação de grupos; g) ponderações; e h) escolha de um mo-

delo estatístico. Embora os estudos considerados envolvam um mesmo assunto, na meta-análise, é importante que grupos homogêneos destes sejam formados e a variabilidade existente entre eles seja modelada e incorporada na análise. No entanto, a formação de grupos homogêneos é uma das maiores dificuldades na meta-análise, podendo ser feita por combinação de níveis de fatores, análises de agrupamento, componentes principais, etc (LOVATTO et al., 2007; MARIANO et al., 2012; NASCIMENTO et al., 2011).

Maiores detalhes sobre meta-análise, podem ser verificados em Fagard, Staessen e Thijs (1996) e Lovatto et al. (2007).

2.3 Redes neurais artificiais

Na tentativa de reproduzir, por meio das máquinas, aspectos do comportamento humano, tais como aprendizado, percepção, raciocínio, evolução e adaptação, surgiu um campo de conhecimento, a Inteligência Artificial. A Inteligência Artificial tem conseguido avanços no emprego de diferentes técnicas que simulam as ações do ser humano na solução de problemas.

Uma das técnicas criadas em Inteligência Artificial, conhecida como Redes Neurais Artificiais (ANN, do inglês *Artificial Neural Networks*), investiga a reprodução do funcionamento e da arquitetura do cérebro em sistemas computacionais, isto é, busca uma forma de representar o conhecimento humano (HAYKIN, 2007).

2.3.1 Neurônio artificial

Buscando modelar o cérebro humano com o uso de pequenos elementos de processamento básicos interconectados (também conhecidos como neurônios artificiais, unidades, nodos ou nós), vários modelos matemático-computacionais

foram desenvolvidos com vários aspectos presentes no neurônio biológico e no comportamento do sistema nervoso, particularmente, os existentes no cérebro humano (HAYKIN, 2007, 2009).

Para entender os conceitos associados às estruturas de uma ANN, é conveniente que se faça uma analogia entre os elementos do neurônio biológico (Figura 1a) e os elementos do neurônio artificial (Figura 1b). Essa analogia é descrita a seguir, com base nos conceitos definidos em Haykin (2007).

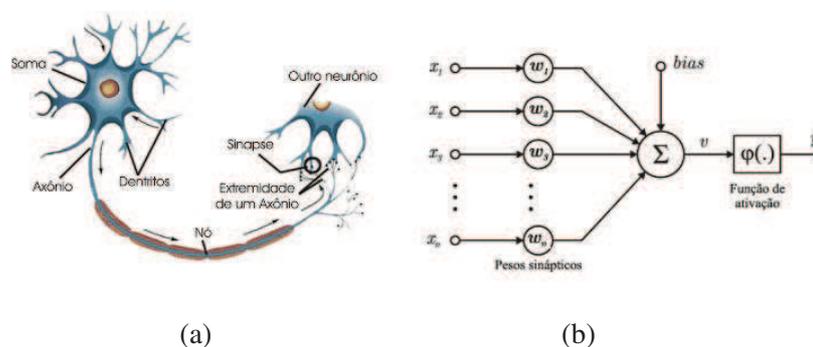


Figura 1 Modelo de neurônio biológico (a) e modelo de neurônio artificial (b)

O funcionamento de um neurônio biológico (Figura 1a) acontece da seguinte maneira: as informações são recebidas pelo neurônio por meio dos dendritos e são processadas no corpo celular. Tal processamento gera um impulso nervoso (estímulo) que é enviado pelo neurônio por meio do axônio. Esse novo estímulo é capturado pelos dendritos dos neurônios vizinhos pelas sinapses. As sinapses são os pontos de contato entre as terminações axônicas de um neurônio e os dendritos do outro, funcionando como válvulas que controlam a transmissão de impulsos (fluxo de informação) entre os neurônios. As sinapses podem ser tanto excitatórias como inibitórias sobre o neurônio receptivo e possuem um papel fundamental na memorização da informação (BRAGA; CARVALHO; LUDERMIR, 2000). Dessa forma, o conjunto das sinapses causa os estímulos no corpo

do neurônio e o somatório de todos os estímulos determina ou não a ativação do mesmo e a transmissão ou não do pulso pelo axônio (BISHOP, 1995; HAYKIN, 2007). E assim, a informação é passada ao neurônio vizinho, recomeçando o ciclo até alcançar o aprendizado.

No modelo artificial (Figura 1b), os três elementos básicos que o compõem são:

- um conjunto de n conexões de entrada (x_1, x_2, \dots, x_n) que correspondem aos dendritos, cujas ligações com o corpo celular artificial são realizadas por meio de elementos chamados de peso (w_1, w_2, \dots, w_n) , simulando as sinapses;
- uma função soma (Σ) , que processa os estímulos captados pelas entradas;
- uma função de ativação $\varphi(\cdot)$, que limita o intervalo do sinal de saída (y) a um valor normalizado. Esse intervalo normalizado da saída de um neurônio é descrito, usualmente, como o intervalo unitário fechado $[0,1]$ ou $[-1,1]$.

Assim, a estrutura de um neurônio artificial considera um conjunto de valores como entradas (x_1, x_2, \dots, x_n) para a produção de uma saída única (y) . Com o intuito de simular o aprendizado, o comportamento das conexões sinápticas entre os neurônios é simulado por meio de seus respectivos pesos (w_1, w_2, \dots, w_n) e somadas ao valor de um *bias* aplicado externamente. O *bias* tem o efeito de aumentar ou diminuir o argumento da função de ativação. É possível considerar o *bias* como uma entrada de valor constante $x_0 = 1$, multiplicado por um peso igual a w_0 . Os valores dos pesos podem ser negativos ou positivos, dependendo se as conexões são inibitórias ou excitatórias. O valor ν , resultante do somatório de todas as entradas ponderadas pelos seus respectivos pesos, é enviado para uma função de ativação $\varphi(\cdot)$, que define a saída (y) do neurônio (FERNEDA, 2006).

Portanto, a saída do neurônio j pode ser escrito matematicamente por:

$$v_j = \sum_{i=0}^n w_{ji}x_{ji} \quad \text{e} \quad y_j = \varphi_j(v_j),$$

em que:

$x_{j1}, x_{j2}, \dots, x_{jn}$ são os sinais de entrada no neurônio j ;

$w_{j1}, w_{j2}, \dots, w_{jn}$ são os pesos sinápticos do neurônio j ;

v_j é a saída do combinador linear em razão dos sinais de entrada do neurônio j , ou seja, o potencial de ativação;

$\varphi_j(\cdot)$ é a função de ativação aplicada no neurônio j e

y_j é a saída do neurônio j .

Outras informações complementares sobre como os neurônios reais inspiraram a criação das redes neurais artificiais podem ser obtidas em Braga, Carvalho e Ludermir (2000), Haykin (2007) e McCulloch e Pitts (1943), além de outros artigos científicos e livros sobre o assunto.

2.3.2 Estrutura e funcionamento de uma ANN

A combinação de vários neurônios artificiais, os quais individualmente têm capacidades relativamente limitadas, compõe uma rede neural artificial, capaz de aprender. Esta rede foi inspirada no que o homem já conhece sobre o cérebro humano, possuindo características de robustez, tolerância a falhas, capacidade de ajuste a novos ambientes pelo treinamento, potencial para processar informações ruidosas, além de alta velocidade de processamento (SEGATTO et al., 2003).

Uma ANN multicamada consiste de uma camada de entrada (input layer), elementos de processamento (neurônios), uma ou mais camadas intermediárias ou ocultas (hidden layers), conexões sinápticas (arestas que ligam uma camada a outra) e uma camada de saída (output layer), conforme Figura 2. Os detalhes de

tal estrutura podem ser verificados em Haykin (2007).

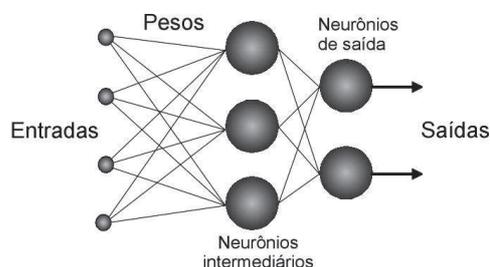


Figura 2 Estrutura de uma rede neural

Conforme a Figura 2, a camada de entrada é composta de várias entradas que simulam a captação de estímulos. Estes estímulos podem ser conectados em muitos outros neurônios da primeira camada intermediária e cada neurônio pode ter várias entradas, mas somente uma saída. Por meio das conexões sinápticas, o sinal de cada saída pode ser utilizado como sinal de entrada a vários neurônios. Da mesma forma, cada neurônio das demais camadas pode receber várias entradas procedentes de outros neurônios. Contudo, cada conexão entre neurônios possui um peso que determina sua contribuição na decisão de disparo (função de ativação), controlando, desta forma, os estímulos.

Todo o processamento supracitado é feito em paralelo, como uma tentativa de reproduzir potencialidades do cérebro humano, em especial a sua capacidade de aprendizagem (HAYKIN, 2009; SEGATTO et al., 2003). Dessa forma, as ANN têm a vantagem de poder “aprender”, errar e aprender com os erros, assim como o cérebro humano. A maior parte desse processamento ocorre nas camadas intermediárias. Contudo, é na camada de saída que a resposta predita é apresentada.

Segundo Haykin (2007), uma rede neural se assemelha ao cérebro humano em dois aspectos:

- o conhecimento é adquirido pela rede, considerando seu ambiente, por um

processo de aprendizagem; e

- forças de conexão entre neurônios, conhecidas como pesos sinápticos, são utilizados para armazenar o conhecimento adquirido.

Uma ANN pode ter centenas ou milhares de unidades de processamento; já o cérebro de um mamífero pode ter muitos bilhões de neurônios (NICHOLLS et al., 2012). Essa diferença impossibilita a total reprodução do comportamento humano com o uso de computadores. O que as máquinas conseguem é desenvolver algoritmos inspirados nos princípios de funcionamento dos neurônios biológicos e na estrutura do cérebro (BISHOP, 1995; HAYKIN, 2007), sendo as mesmas possíveis de serem implementadas tanto em hardware quanto em software. A palavra inspirado é usada para deixar claro que o modelo matemático é muito simples quando comparado com a complexidade de um neurônio real (COSTA, 2009).

Embora simples, as ANN possuem a capacidade de executar uma série de tarefas como aproximação de funções e predição, controle, filtragem não linear, reconhecimento e classificação de padrões, *clustering* ou agrupamento, previsão de séries temporais, otimização, processamento de sinais, análise de imagens e controle de processos. Algumas aplicações de ANN usadas como predição podem ser encontradas em Ahmadi et al. (2008), Ahmadi, Mottaghitlab e Nariman-Zadeh (2007), Gheyas e Smith (2011), Okut et al. (2011) e Perai et al. (2010).

2.3.3 Um pouco de história

Em 1958, Frank Roseblatt mostrou o modelo do *perceptron* (neurônio), resultando em um livro publicado em 1962. O *Perceptron* de Roseblatt consistia em um sistema de classificação de padrões, organizados em camada de entrada e saída (apenas uma camada de neurônios). Nele, os pesos das conexões eram adaptados, a fim de se atingir a eficiência sináptica, usada no reconhecimento de

caracteres, no aprendizado de funções lógicas. Os neurônios do *perceptron* são similares ao de McCulloch-Pitts (MCCULLOCH; PITTS, 1943) por terem a função de ativação do tipo degrau, mas possuem pesos associados e *bias*.

Em 1969, Marvin Mjinsky e Seymour Papert publicaram um livro chamado “*Perceptrons*”, constatando que um neurônio do tipo *perceptron* só é capaz de resolver problemas com dados de classes linearmente separáveis, ou seja, que podem ser separados por uma reta em um hiperplano (Figura 3a). Isto provocou uma parada e um grande atraso no desenvolvimento de ANN.

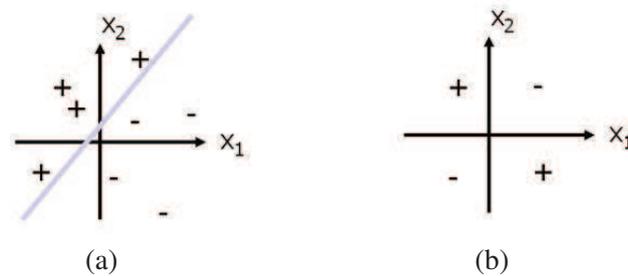


Figura 3 Exemplos linearmente separável (a) e não separável linearmente (b)

Nos anos 80, frente ao aumento dos recursos computacionais disponíveis houve um reaquecimento das pesquisas em redes neurais. Mas foi só em 1986, que James McClelland e David Rumelhart apresentaram o algoritmo *backpropagation* para o treinamento de um *perceptron* multicamadas, que é uma generalização do *perceptron* simples. Assim, para resolver problemas complexos que envolvem padrões não separáveis linearmente (Figura 3b) devem ser usadas redes *perceptron* multicamadas (HAYKIN, 2009).

2.3.4 *Perceptron* multicamadas - MLP

O *perceptron* multicamadas, conhecido no inglês como *Multilayer Perceptron* (MLP), é o modelo de rede neural artificial mais conhecido e mais utilizado.

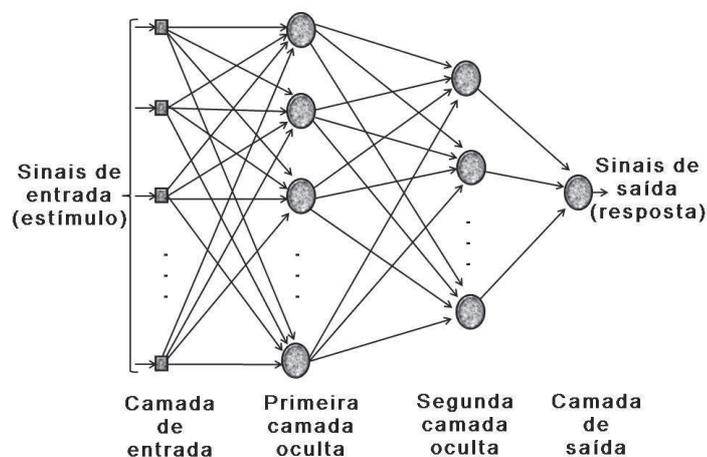


Figura 4 Estrutura de um perceptron multicamadas com duas camadas intermediárias

Este modelo é um tipo de rede direta (*feedforward*), em que os neurônios são dispostos em camadas e as informações percorrem apenas o sentido da entrada para a saída (Figura 4). Esta seção descreve o que é uma rede *MLP*, baseada em Bishop (1995) e Haykin (2007).

Em suma, os sinais de entrada são apresentados à camada de entrada. Em seguida, essas informações são repassadas para os neurônios da camada intermediária, que irão processar a entrada. Se a rede possui mais de uma camada intermediária, as saídas de cada camada são reprocessadas e enviadas para a próxima camada. Quando chegar à última camada, então será realizado o último processamento, que resultará em um vetor de saídas de tamanho igual ao número de neurônios da última camada (Figura 4).

Segundo Cybenko (1988), uma rede *MLP* que contenha uma camada intermediária pode aproximar qualquer função matemática contínua e, se houver duas camadas ocultas, a rede pode aproximar qualquer função contínua ou não contínua.

Portanto, redes *MLP* são consideradas aproximadores universais de funções, pois não são necessárias mais que duas camadas ocultas para resolver um problema de aproximação e predição.

Para se obter os pesos sinápticos que interconectam os neurônios nas diferentes camadas de uma rede *MLP*, é comum utilizar algoritmos de treinamento supervisionado. Nestes, um supervisor externo apresenta à rede neural alguns conjuntos de padrões de entrada e seus correspondentes desejados padrões de saída (valores observados). Portanto, é necessário ter um conhecimento prévio do comportamento que se deseja ou se espera da rede. À medida que se obtêm os valores preditos pela rede, a mesma é comparada com os valores observados e o supervisor informa o erro verificado à rede, para que sejam feitos ajustes nos pesos, a fim de melhorar suas futuras respostas. Um dos algoritmos mais difundidos para treinamento supervisionado é o algoritmo retropropagação, conhecido como *back-propagation*, no inglês (RUMELHART; MCCLELLAND, 1986). Uma descrição detalhada pode ser encontrada na seção 2.3.7.

2.3.5 Parâmetros de treinamento de uma ANN

Alguns parâmetros que devem ser pré-estabelecidos antes da implementação de uma ANN e que, portanto, estão sujeitos a modificações são:

- Topologia das interconexões
É a forma como as camadas são conectadas (redes recorrentes ou redes não-recorrentes).
- Número de camadas (*layers*)
Quanto maior o número de camadas, maior será a complexidade e o tempo de processamento da rede.

- Número de neurônios por camada

Também determina a complexidade do modelo de ANN escolhido. O número de neurônios na camada de entrada da rede é determinado pela dimensionalidade do espaço de observação. Já na camada de saída, esse número depende da dimensão requerida pelo vetor de saídas. Contudo, para determinar exatamente o número de neurônios necessários na camada intermediária, testes devem ser feitos treinando diferentes arquiteturas e comparando suas saídas, buscando aquela que garante uma melhor generalização.

- Função de ativação

A seleção das funções de ativação é uma etapa importante na configuração de uma ANN e são, usualmente, consideradas idênticas para todos os neurônios de uma mesma camada. A função de ativação pode apresentar características lineares ou não-lineares, determinando a saída de um neurônio a partir do seu potencial de ativação (v_k). Em geral, a simulação de um neurônio biológico é mais precisa com o uso de uma função não-linear. A Figura 5 ilustra o comportamento das funções de ativação mais difundidas na literatura, sendo a função sigmoide uma das mais utilizadas. Além dessas funções, existem outras e algumas variações destas que, também, são usadas como função de ativação.

- Número de épocas

O treinamento de uma rede é realizado por épocas, em que uma época corresponde a cada vez que os vetores de entrada-saída do conjunto de treinamento é apresentada à rede.

- Taxa de aprendizagem

A taxa de aprendizagem indica a velocidade com que o vetor de peso será

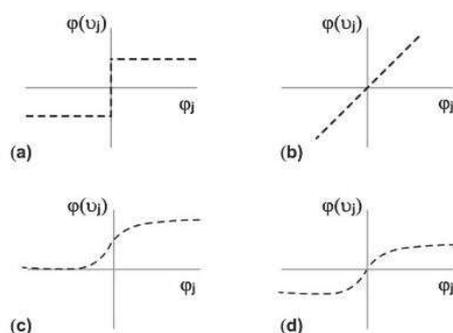


Figura 5 Tipos de função de ativação: (a) função degrau, (b) função linear, (c) função sigmoide, (d) função tangente hiperbólica

atualizado. Ela pode variar de 0 a 1, porém é comum utilizar valores entre 0,2 e 0,8. Uma taxa de aprendizagem muito baixa torna o aprendizado da rede muito lento. Já uma taxa de aprendizagem muito alta provoca oscilações e impede a convergência no processo de treinamento.

- Taxa de momento

A taxa de momento é adicionada ao cálculo de atualização dos pesos do algoritmo de treinamento. Ela acelera o tempo de aprendizagem durante o treinamento e estabiliza a convergência. Além de ser aconselhada para diminuir o tempo de treinamento, a taxa de momento, também, é útil para solucionar o problema dos mínimos locais, possibilitando o algoritmo livrar-se de mínimos locais durante o processo de busca pelo mínimo global (ponto de menor erro) (THOMÉ, 2012). Essa taxa pode variar de 0 a 1, sendo costume utilizar valores entre 0,1 e 0,5. Afinal, altas taxas de momento podem fazer com que o algoritmo de treinamento passe pelo mínimo global, prejudicando o treinamento.

- Especificação dos pesos sinápticos

Esta especificação envolve a utilização de algum algoritmo de treinamento.

2.3.6 Treinamento e teste

Um dos fatores mais importantes de uma ANN é a sua capacidade de generalização, baseada na habilidade que ela tem de aprender por meio de exemplos. Isso é feito por meio de um processo iterativo de ajustes aplicado a seus pesos, chamado de aprendizagem ou treinamento. Como já foi mencionado anteriormente, são estes pesos os responsáveis por armazenar e processar o conhecimento.

De posse do conjunto de dados, grande parte destes devem ser escolhidos aleatoriamente e separados para o treinamento da rede. Este conjunto de dados é conhecido como dados de treinamento. O restante dos dados, conhecido como dados de teste, serão utilizados na fase de teste da rede. É importante salientar que esses dois conjuntos devem ser disjuntos; os dados de treinamento não devem ser utilizados para testar a rede, porque a rede pode memorizar o padrão de dados e o que se espera é que a rede “aprenda” as regras e não “decore” os exemplos.

Assim, o treinamento da rede inicia-se com a apresentação dos dados de treinamento à rede e termina quando suas conexões (pesos) estão devidamente treinadas, ou seja, o aprendizado da rede nada mais é do que a busca pelos pesos corretos e ocorre quando a rede neural atinge uma solução generalizada para uma classe de problemas. Durante esse treinamento, a rede incorpora informações com base em exemplos que lhe são apresentados.

O treinamento deve ser interrompido quando a rede apresentar uma boa capacidade de generalização e quando a taxa de erro for suficientemente pequena, ou seja, menor que um erro pré-estabelecido. Assim, deve-se encontrar um ponto ótimo de parada com erro mínimo e capacidade de generalização máxima. Caso contrário, a rede pode se especializar no conjunto de dados do treinamento, causando problema de *overfitting*. Para evitar esse problema, recomenda-se considerar algumas técnicas a fim de otimizar a generalização das redes como comparação

dos modelos, critérios de parada, validação cruzada ou *pruning* (HAYKIN, 2007).

No processo de treinamento, são utilizados alguns, ou pelo menos um, dos seguintes critérios de parada:

- Erro mínimo: o treinamento é encerrado quando o algoritmo de treinamento levar a rede a convergir para um erro menor que o mínimo estipulado como critério de convergência (THOMÉ, 2012).
- Número de épocas: o treinamento é encerrado quando se atinge um número máximo de vezes pré-estabelecido para que o conjunto de treinamento seja apresentado à rede. A determinação do número ideal de épocas de treinamento da rede, também, é realizada por meio de tentativa e erro. Um número excessivo de épocas pode levar a rede à perda do poder de generalização (*overfitting*), pois a rede entra em um processo de “memorização” dos padrões de entrada. Porém, se um número muito pequeno for pré-estabelecido, a rede torna-se incapaz de representar os dados (BASHEER; HAJMEER, 2000).
- Critério da parada antecipada, do inglês *early-stopping*: é uma técnica de determinação do número de épocas de treinamento que garante uma boa capacidade de generalização da rede. Nesta técnica, é possível identificar o início do excesso de treinamento, pelo uso da validação cruzada. Na validação cruzada, os dados utilizados para o treinamento da rede são divididos, aleatoriamente, em dois conjuntos distintos, conjuntos de treinamento e de validação. É importante que estes dois conjuntos possuam uma boa representatividade do problema. Assim, a cada época do treinamento, os dados de validação são apresentados à rede e o erro de validação é calculado. Quando finalizada essa fase de validação, o treinamento é reiniciado

em uma nova época. Dessa forma, o processo de treinamento da rede é interrompido quando o erro do conjunto de validação deixa de ser monotonicamente decrescente por algumas épocas (HAYKIN, 2007).

A Figura 6 mostra os erros para os conjuntos de treinamento e de validação. Observa-se que a curva de aprendizagem do conjunto de validação decresce para um mínimo e, então, começa a crescer enquanto a curva de treinamento continua decrescendo. Dessa forma, o ponto ótimo da parada do treinamento ocorre no menor valor de erro quadrático médio (m.s.e.) para o conjunto de validação, garantindo uma boa generalização e evitando a memorização dos dados.

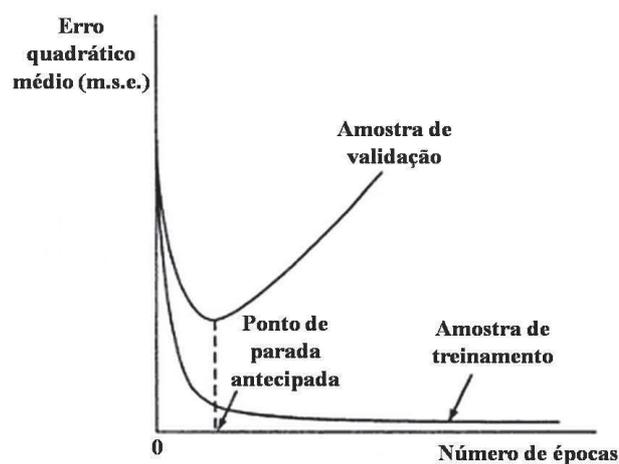


Figura 6 Regra da parada antecipada baseada na validação cruzada

A partir da época determinada como ponto ótimo de parada do treinamento, a rede começa um processo de memorização dos dados de treinamento, apresentando um grande erro de validação e, assim, a rede não será capaz de determinar bons resultados quando usada com dados desconhecidos.

Após o treinamento, o conjunto de teste é utilizado para averiguar a capacidade de generalização da rede, ou seja, verificar o desempenho de previsão da ANN. Este processo recebe o nome de teste da rede. Se a rede foi bem treinada, os dados de teste disponibilizarão um resultado final coerente, indicando que a rede tem uma boa performance real. Finalmente, com a rede treinada e testada, ela pode ser disponibilizada para o uso com novos dados de entrada (BRAGA; CARVALHO; LUDERMIR, 2000; HAYKIN, 2007).

Uma consideração importante é usar o maior conjunto de dados possível no desenvolvimento da rede neural, porque isso melhora a possibilidade de se obter respostas mais acuradas. Além disso, as estatísticas usadas para verificar a convergência das redes são, geralmente, melhores estimadas quando o conjunto de dados é maior (BISHOP, 1995).

2.3.7 Algoritmos de treinamento

O procedimento utilizado para realizar o processo de aprendizagem é chamado de algoritmo de treinamento ou de aprendizagem. A função deste algoritmo é modificar os pesos da rede de uma forma ordenada, para que a rede consiga “aprender” ao longo do treinamento e tomar decisões apropriadas (HAYKIN, 2007). A seguir serão descritos dois algoritmos: o *Backpropagation* e o Levenberg-Marquardt.

Backpropagation

Este algoritmo é conhecido como algoritmo de retropropagação do erro (*error backpropagation*), o qual é baseado na regra de aprendizagem por correção do erro (RUMELHART; MCCLELLAND, 1986).

A aprendizagem por retropropagação do erro ocorre em dois sentidos:

passo para frente, a propagação (*forward*), e passo para trás, a retropropagação (*backward*). No passo para frente, o padrão de entrada (vetor de entrada) se propaga pela rede, camada por camada, gerando um vetor de saídas (respostas) calculadas pela rede. Durante a propagação, os pesos da rede são todos fixos (HAYKIN, 2007) e, normalmente, os pesos iniciais são adotados como números aleatórios. Já no passo para trás, os pesos são todos ajustados de acordo com uma regra de correção de erros. O sinal do erro é, então, propagado para trás pela rede, contra a direção das conexões sinápticas, com o intuito de que os pesos sejam ajustados e atualizados (HAYKIN, 2007; SEGATTO et al., 2003).

O algoritmo do *backpropagation* é, então, realizado da forma descrita a seguir (CAMPOS et al., 2010; HAYKIN, 2001; RUMELHART; MCCLELLAND, 1986).

Um vetor padrão \mathbf{X} é apresentado à rede, a qual fornece uma saída \mathbf{Y} . O erro na saída do neurônio j , da camada de saída, é calculado para cada iteração k (isto é, a apresentação do k -ésimo vetor de entrada dos dados de treinamento). Logo, o valor instantâneo da soma dos erros quadráticos dos neurônios da camada de saída é definido como:

$$\varepsilon(k) = \frac{1}{2} \sum_{j \in C} (e_j(k))^2 = \frac{1}{2} \sum_{j \in C} (d_j(k) - y_j(k))^2,$$

em que o conjunto C inclui todos os neurônios da camada de saída, $d_j(k)$ é a saída observada (desejada) do neurônio j da camada de saída na iteração k e $y_j(k)$ é a saída calculada pela rede do mesmo neurônio j da mesma iteração k .

A média dos valores instantâneos $\varepsilon(k)$ para todos os n vetores de entrada (amostras) do conjunto de treinamento, correspondente ao erro quadrático médio

(*m.s.e.*) de uma época, é dada por:

$$m.s.e. = \frac{1}{n} \sum_{k=1}^n \varepsilon(k).$$

O objetivo do processo de aprendizagem é, por um método iterativo de gradiente, ajustar os parâmetros livres (pesos e bias) da rede, de forma a minimizar o *m.s.e.* Busca-se encontrar uma relação entre os vetores de entradas e saídas fornecidos, de tal forma que as saídas calculadas pela rede se aproximem das saídas desejadas. Assim, o ajuste dos pesos é feito de acordo com os respectivos erros calculados para cada padrão apresentado à rede (HAYKIN, 2007).

Seja

$$v_j(k) = \sum_{i=0}^m w_{ji}(k)y_i(k),$$

em que $v_j(k)$ é o potencial de ativação aplicado na entrada da não-linearidade associada ao neurônio j da iteração k ; m é o número total de entradas aplicadas ao neurônio j da iteração k ; $y_i(k)$ é o sinal da i -ésima entrada do neurônio j da iteração k (= sinal na saída do neurônio i , pertencente à camada esquerda da que pertence o neurônio j , se este não pertencer à primeira camada escondida; caso o neurônio j pertença à primeira camada escondida, então $y_i(k)$ corresponde ao i -ésimo nó de entrada $x_i(k)$ do *MLP*) e $w_{ji}(k)$ é o peso sináptico conectando a saída do neurônio i ao neurônio j .

Dessa forma, o sinal $y_j(k)$ resultante na saída do neurônio j na iteração k é dado por:

$$y_j(k) = \varphi_j(v_j(k)).$$

Considerando-se o neurônio de índice j da rede, e utilizando-se o método do gradiente descendente, o ajuste de pesos pode ser descrito, associado à derivada

parcial do erro quadrático, da seguinte forma:

$$w_{ji}(k+1) = w_{ji}(k) + \Delta w_{ji}(k),$$

sendo

$$\Delta w_{ji}(k) = -\eta \frac{\partial \varepsilon(k)}{\partial w_{ji}(k)}. \quad (2.1)$$

Na Eq. 2.1, conhecida como regra delta, $\frac{\partial \varepsilon(k)}{\partial w_{ji}(k)}$ é a derivada parcial do erro em relação ao peso w_{ji} e η é a constante que determina a razão de aprendizado do algoritmo *backpropagation*, ou seja, é a taxa de aprendizado. Na regra delta, o uso do sinal negativo mostra qual a direção a ser adotada para minimizar a função do erro quadrático correspondente, isto é, deve-se adotar a direção contrária ao gradiente. Outra forma de representar essa regra é:

$$\Delta w_{ji}(k) = \eta \delta_j(k) y_i(k),$$

em que $\delta_j(k)$ é o gradiente local do neurônio j na iteração k , definido por

$$\delta_j(k) = \varphi'_j(v_j(k)) e_j(k).$$

Esta última equação demonstra que o gradiente local $\delta_j(k)$ para o neurônio j da camada de saída é igual ao produto do respectivo sinal de erro pela derivada da função de ativação relacionada.

Já para o caso de um neurônio j localizado em uma camada intermediária, o sinal de erro deve ser determinado de forma recursiva, em termos dos sinais de erro de todos os neurônios aos quais o neurônio oculto está diretamente conectado.

Dessa forma, o cálculo do gradiente local é redefinido para:

$$\delta_j(k) = \varphi'_j(v_j(k)) \sum_{p \in R(j)} \delta_p(k) w_{pj}(k),$$

em que $R(j)$ é o conjunto dos índices dos neurônios que se encontram na camada à direita do neurônio j e que estão interligados a este neurônio j .

O treinamento das redes *MLP* com *backpropagation* pode demandar muitos passos no conjunto de treinamento, resultando em uma convergência lenta ou até na estacionaridade do erro em mínimos locais. Assim, quanto maior for a taxa de aprendizagem, maior será a mudança nos pesos, aumentando a velocidade do aprendizado; mas isso pode levar a uma oscilação do modelo na superfície de erro. O ideal seria utilizar a maior taxa de aprendizado possível que não levasse a uma oscilação, resultando em um aprendizado mais rápido. Uma maneira de aumentar a taxa de aprendizagem η , evitando causar a instabilidade da rede, é a inclusão da taxa de momento α na regra delta generalizada, fazendo:

$$\Delta w_{ji}(k) = \alpha \Delta w_{ji}(k-1) + \eta \delta_j(k) y_i(k).$$

Em cada iteração, feito o ajuste dos pesos, um novo padrão é apresentado à rede e este processo é repetido até que o treinamento da rede seja interrompido pelo critério de parada estabelecido. Este é o modo de treinamento padrão-a-padrão ou sequencial (*pattern mode*), em que os pesos são atualizados após a apresentação de cada padrão à rede neural. Outro modo de treinamento deste algoritmo é o acumulativo ou por lote (*batch mode*), em que os pesos são ajustados apenas depois da apresentação de todos os padrões, ou seja, após cada época. Muitas variações do algoritmo básico utilizam essa forma de aprendizado como tentativa de acelerar o processo de treinamento da rede (HAYKIN, 2001).

Neste algoritmo, as funções de ativação podem ser não-linearizáveis, desde que sejam diferenciáveis (RUMELHART; MCCLELLAND, 1986). Uma função de ativação amplamente utilizada é a função sigmoide. Para maiores detalhes sobre este algoritmo, recomenda-se a leitura do artigo original de Rumelhart e McClelland (1986).

Levenberg-Marquardt

O algoritmo *backpropagation* padrão utiliza o método do gradiente, no decorrer do processo de aprendizagem, de forma a minimizar a função erro. Já o algoritmo de Levenberg-Marquardt (LM) utiliza uma aproximação pelo método de Newton, que faz uso da matriz Hessiana ($\mathbf{H} = \mathbf{J}^t \mathbf{J}$). No entanto, nesta aproximação é introduzido um parâmetro μ .

Conforme Bishop (1995), a atualização dos pesos pelo algoritmo LM é dado por:

$$\mathbf{w}(k+1) = \mathbf{w}(k) - (\mathbf{J}^t(k)\mathbf{J}(k) + \mu_k \mathbf{I})^{-1} \mathbf{J}^t(k) \mathbf{e}(k)$$

em que \mathbf{I} é a matriz identidade, $\mathbf{w}(k)$ é o vetor de pesos na época k ; $\mathbf{e}(k) = (e_1(k), e_2(k), \dots, e_n(k))$ é o vetor de erros entre os valores observados e preditos das n amostras de treinamento, na época k ; $\mathbf{J}(k)$ é a matriz Jacobiana do vetor de erros $\mathbf{e}(k)$ em relação ao vetor de pesos $\mathbf{w}(k)$ na época k e μ_k é a constante do método de LM na época k .

O parâmetro μ consiste em 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 LM apresenta convergência em menos iterações (no caso, épocas), mas requer mais cálculos por iteração devido ao cálculo de matrizes inversas. Por isso, é considerado o método mais rápido para treinamento de redes *feedforward*,

que possui uma quantidade moderada de pesos sinápticos. Se esse número de pesos é elevado, a utilização desse algoritmo é pouco prática (BARBOSA; FREITAS; NEVES, 2005).

Conforme relatado por Brito Júnior et al. (2007), se μ for igual a zero, este é somente o método de Newton, usando a matriz Hessiana aproximada. O método de Newton é mais rápido e preciso ao aproximar-se do erro mínimo. Assim, μ é diminuído após cada etapa bem sucedida e é aumentado se, após uma nova tentativa, ocorrer um aumento da discrepância entre a saída da rede e o alvo a ser considerado. Baseado neste procedimento, a discrepância sempre será reduzida a cada nova iteração do algoritmo.

No entanto, se μ tornar-se muito grande, este algoritmo sofrerá instabilidades numéricas. Para contornar esse problema, Marquardt propôs substituir a matriz identidade pela matriz diagonal de \mathbf{H} ($\text{diag}(\mathbf{H})$), fazendo com que cada componente do gradiente seja ponderado de acordo com sua curvatura. Dessa forma, existirá uma tendência de convergência na direção em que o gradiente é menor (FRANÇA et al., 2009).

2.3.8 Vantagens e desvantagens das ANN

Redes neurais podem ser descritas como uma estratégia para a modelagem matemática de problemas, que são concebidos como sistemas com entradas e saídas. Conforme descrito em Balcean e Ooghe (2004), dentre as vantagens das redes neurais na aproximação de funções e predição, destacam-se: sua capacidade de aprender soluções gerais com base nos dados fornecidos para o treinamento (aprendizagem adaptativa), sua habilidade de analisar dados complexos, não é preciso considerar hipóteses restritivas, não é preciso conhecer a relação matemática entre as entradas e saídas e, também, podem ser utilizadas inúmeras variáveis de

entradas que possam interferir na variabilidade da variável predita.

Embora as ANN sejam consideradas como robustas e flexíveis, existem algumas desvantagens no uso das mesmas: são requeridos dados com alta representatividade, há risco de *overfitting*, a arquitetura da rede deve ser cuidadosamente pré-definida, é necessária uma grande amostra de treinamento e a rede é considerada como uma “caixa preta”, pois não se tem uma visão total do funcionamento interno da ANN. Além disso, a rede não fornece estimativas de parâmetros, como em uma regressão. Para esta última desvantagem, Yee, Prior e Florence (1993) defendem que, embora seja sempre comum interpretar o significado biológico das estimativas dos parâmetros de uma equação, na realidade, pode ser mais prático ignorar a relevância das estimativas dos parâmetros e centrar-se na capacidade de prever respostas.

2.4 Comitê de redes neurais

Um modelo de rede neural nem sempre é considerado totalmente satisfatório para um problema específico, pois seu desempenho depende do ajuste de muitos parâmetros de treinamento, além do grau de dificuldade associado ao problema. As várias redes implementadas com diferentes parâmetros de treinamento, são testadas e aquela com melhor desempenho (menor m.s.e.) é escolhida como a rede com melhor generalização. No entanto, a seleção de uma só rede acaba descartando informações relevantes. Por isso, para a obtenção de resultados mais acurados em modelos de ANN, é recomendável a utilização de um comitê de redes.

Um comitê de redes, o qual é baseado no princípio de dividir e conquistar (HAYKIN, 2007) consiste em um método de aprendizado supervisionado ou não-supervisionado, cujo objetivo é aumentar a capacidade de generalização de estimadores, evitando o sobreajuste dos dados e a convergência da rede para um

mínimo local.

Como o próprio nome indica, um comitê de redes agrega, de alguma forma, o conhecimento adquirido por mais de um componente para se chegar a uma solução única e global que é supostamente superior àquela obtida por qualquer um dos componentes isolados (HAYKIN, 2007; LIMA, 2004). Cada componente isolado é um modelo de ANN com capacidade de generalização para determinado problema, denominado como rede especialista. Assim, as redes especialistas que apresentam melhor capacidade de generalizar são selecionadas para compor o comitê.

O primeiro trabalho que consta na literatura sobre o uso de um comitê de redes foi o de Nilsson (1965), em que ele considerou a estrutura de uma rede neural composta de uma camada de *perceptrons* elementares, seguida de um *perceptron* responsável por realizar um esquema de votação na segunda camada. Desde então, em virtude de sua praticidade, eficácia na generalização do aprendizado e baixo custo computacional extra, essa metodologia vem sendo aplicada em diversas áreas de interesse científico e tecnológico (KENARI; MASHOHOR, 2013; LIPNICKAS, 2008; OLIVEIRA et al., 2010), apresentando resultados mais acurados que qualquer um dos componentes isolados.

Segundo Haykin (2007), comitês de redes são aproximadores universais e podem se apresentar em estruturas dinâmicas ou estáticas. Nas estruturas dinâmicas, o sinal de entrada faz parte do mecanismo que integra as saídas dos especialistas individuais em uma saída global por meio de uma rede de passagem. São conhecidos como mistura de especialistas, onde as saídas dos diferentes especialistas são não-linearmente combinados por uma rede única, ou como mistura hierárquica de especialistas, onde as saídas dos diferentes especialistas são não-linearmente combinados por várias redes organizadas de forma hierárquica.

Nas estruturas estáticas, as redes especialistas são treinadas de modo a compartilhar uma entrada comum e suas saídas individuais são combinadas, por meio de um mecanismo que não envolve o sinal de entrada, para se produzir uma saída única. Essa categoria inclui: a média de *ensemble*, reforço e votação.

- Geração dos vários componentes

Os principais requisitos que cada um dos componentes de um comitê deve apresentar são: um bom desempenho quando tomado isoladamente e uma generalização de forma não-correlacionada frente aos demais, isto é, que não apresente os mesmos erros de predição. Segundo Bishop (1995), comitês de redes podem ser mais precisos que as redes individuais que compõem o comitê, desde que as redes apresentem diferenças na generalização. Conforme Lima (2004) relatou, há vários métodos utilizados durante a geração de componentes que buscam maximizar a divergência de comportamento, ao mesmo tempo em que se busca garantir bons desempenhos. Estes métodos podem ser divididos em duas classes:

1. geração pelo pré-processamento de parâmetros e aspectos estruturais de redes neurais artificiais, incluindo a manipulação de parâmetros como condição inicial (MACLIN; SHAVLIK, 1995), algoritmo de treinamento (KENARI; MASHOHOR, 2013), dados de treinamento e arquitetura da rede (OLIVEIRA et al., 2010).
2. geração pelo pré-processamento dos dados de treinamento, incluindo reamostragem dos dados (*bagging*) (KORDOS et al., 2011), reamostragem adaptativa (*boosting*) (RUSTEMPASIC; CAN, 2013), conjunto de treinamento disjuntos, variáveis distintas e transformação de variáveis (TSYMBAL; PECHENIZKIY; CUNNINGHAM, 2005).

Maiores detalhes podem ser vistos em Haykin (2007) e Lima (2004).

- Métodos de combinação

Após a escolha do conjunto de redes especialistas a compor o comitê, o passo seguinte é a escolha dos métodos para combinação de suas saídas. Há vários métodos diferentes de combinação específicos para problemas de classificação e regressão. Algumas técnicas de composição de comitês mais utilizadas são: voto majoritário (mais que a metade concorda), média, média ponderada, aproximação Bayesiana, agregação probabilística, agregação por redes neurais, unanimidade (todos aceitam), unanimidade modificada (todos aceitam ou rejeitam), entre outras (BACAUSKIENE; VERIKAS, 2004; CANUTO et al., 2007; LIMA, 2004).

Na média de *ensemble* (*ensemble averaging*), isto é, utilizando o estimador da média para agregar as redes do comitê, as respectivas saídas dos diferentes componentes são combinadas linearmente para produzir uma saída global. Muitos trabalhos utilizaram a média para resumir a informação gerada dos vários componentes dos comitês de redes, obtendo resultados satisfatórios (LIMA, 2004; LIPNICKAS, 2008; SIWEK; OSOWSKI; SZUPILUK, 2009). Apesar de ser um dos métodos mais simples utilizado em um problema de regressão, Bishop (1995) mostrou que um comitê de redes por estimador da média terá um erro esperado menor que a média dos erros individuais de cada especialista considerado no comitê.

A Figura 7 ilustra a estrutura de um comitê por *ensemble*, o qual combina a saída de N redes especialistas, sendo os especialistas identificados por MLP_i , com $i = 1, \dots, N$.

Na Figura 7, a primeira camada consiste de redes especialistas individuais, que recebem os mesmos dados de entrada, podem ou não ter diferentes parâmetros de treinamento, mas os vetores de saída com mesma dimensão. A segunda camada

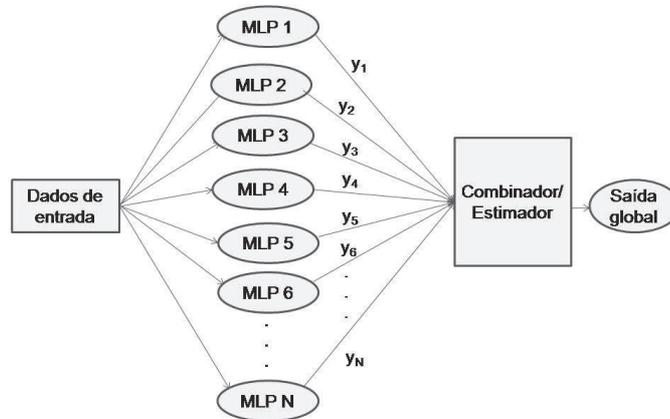


Figura 7 Estrutura estática de um comitê com N redes neurais

consiste num processo de combinação que opera nas saídas da camada anterior para gerar o resultado (saída) global do comitê.

Algumas vantagens no uso de um comitê de redes é que o mesmo evita a perda de informação e o risco de *overfitting*, corrige o problema de mínimos locais e, usualmente, produz melhor resultado que um único componente. Além disso, o tempo de treinamento de uma única rede neural complexa (muitos neurônios), provavelmente, é maior do que o caso em que comitês com especialistas mais simples são treinadas em paralelo.

2.5 Critérios de adequabilidade de ajuste

Ao se implementar uma ANN, é necessário estabelecer algum critério que permita a avaliação da aprendizagem da rede neural. A medida mais comum de desempenho encontrada na bibliografia é a taxa de erro. Este erro pode ser computado de várias formas e as medidas mais utilizadas na verificação da adequabilidade de ajuste de modelos de redes neurais são: R^2 , MAD , $MAPE$, $Bias$ e $m.s.e.$

(BOLZAN; MACHADO; PIAIA, 2008; PERAI et al., 2010). Essas estatísticas indicam as medições de erro de predição, com base na diferença entre os valores observados e preditos e podem ser calculados pelas seguintes expressões:

- R^2 - coeficiente de determinação

$$R^2 = 1 - \frac{\sum_{i=1}^n (d_i - \hat{y}_i)^2}{\sum_{i=1}^n (d_i - \bar{d}_i)^2}$$

- MAD - desvio médio absoluto

$$MAD = \frac{\sum_{i=1}^n |d_i - \hat{y}_i|}{n}$$

- $MAPE$ - erro médio percentual absoluto

$$MAPE = \frac{\sum_{i=1}^n \left| \frac{d_i - \hat{y}_i}{d_i} \right|}{n} * 100,$$

com $d_i \neq 0$.

- $Bias$ - viés (erro médio)

$$Bias = \frac{\sum_{i=1}^n (d_i - \hat{y}_i)}{n}$$

- m.s.e. - erro quadrático médio

$$m.s.e. = \frac{\sum_{i=1}^n e_i^2}{n} = \frac{\sum_{i=1}^n (d_i - \hat{y}_i)^2}{n}$$

em que d_i é o i -ésimo valor observado (ou desejado), \hat{y}_i é o i -ésimo valor predito, n é o número de valores observados e \bar{d}_i é a média dos n valores observados.

Essas medidas de adequabilidade de modelos são utilizadas na comparação e seleção de modelos. O *Bias* indica o quanto o modelo foi superestimado ou subestimado. Portanto, aqueles modelos com alto valor de R^2 , valor de *Bias* mais próximo a zero e outras medidas de erro (*MAD*, *MAPE*, *m.s.e.*) com baixos valores são desejados e indicam um modelo mais acurado.

Se a diferença entre o valor observado e o valor predito é pequena (menor que um), quando ele é elevado ao quadrado tem-se um erro menor ainda. Em contrapartida, se um elemento apresenta um erro alto (maior que um), quando este é elevado ao quadrado o valor do erro aumenta. Portanto, considerando o *m.s.e.* é possível identificar mais rapidamente quando o treinamento não está adequado, já que a taxa de erro pode aumentar drasticamente.

2.6 Porcentagem de acerto da predição - PCP

A porcentagem de acerto da predição (percentage of correct predictions - PCP) é estabelecida, neste trabalho, com o objetivo de determinar qual a taxa de acerto do modelo considerado. Esta medida é, também, utilizada na comparação da capacidade de predição de modelos. Para a aplicação da PCP, são construídos intervalos de predição para cada amostra do conjunto de teste. Estes intervalos levam em consideração alguma incerteza sobre as estimativas de interesse. Dessa

forma, é possível verificar se o modelo considerado garante predições acuradas em relação aos valores observados (reais). Se o valor observado estiver incluído no intervalo de predição de determinada amostra, não há nenhuma evidência de que o valor observado difere do valor predito pelo modelo.

A PCP é obtida considerando a frequência dos valores observados das amostras, do conjunto de teste, dentro dos respectivos intervalos de predição. Esta porcentagem é definida por

$$PCP = \frac{\sum_{i=1}^K I(LI_i \leq Y_{OBS_i} \leq LS_i)}{K} \cdot 100,$$

em que:

$I(.)$ é uma função indicadora, com $I(.) = \begin{cases} 1, & \text{se } Y_{OBS_i} \in \text{intervalo} \\ 0, & \text{caso contrário} \end{cases}$;

LI_i e LS_i são, respectivamente, os limites inferior e superior do intervalo calculado para a i -ésima amostra do conjunto de teste (no presente trabalho, o i -ésimo alimento);

Y_{OBS_i} é o valor observado (valor real de EMAn) para a i -ésima amostra;

K é o número de amostras contidas no conjunto de teste.

2.7 Intervalos de máxima densidade de probabilidade - HPD

Os intervalos de credibilidade são os correspondentes bayesianos aos intervalos de confiança na abordagem clássica. Considerando θ uma quantidade de interesse desconhecida ou aleatória, intervalo de credibilidade bayesiano é um intervalo C de credibilidade de $100(1 - \alpha)\%$, ou nível de credibilidade $1 - \alpha$, para θ se $P(\theta \in C) \geq 1 - \alpha$.

Dessa forma, a definição probabilística expressa a pertinência de θ pertenc-

cer ao intervalo C . Porém, com o mesmo nível de credibilidade é possível construir uma infinidade de intervalos usando a definição acima. Carvalho (2009) descreve três formas de construir um intervalo de credibilidade (Figura 8):

- (a) selecionar o menor intervalo possível, o qual, para uma distribuição unimodal envolverá a escolha dos valores de mais alta densidade de probabilidade ($P(\theta < A) + P(\theta > B) = \alpha$);
- (b) selecionar o intervalo baseado nos quantis, em que a probabilidade de estar abaixo do limite inferior é a mesma de estar acima do limite superior ($P(\theta < A) = P(\theta > B) = \alpha/2$);
- (c) selecionar o intervalo no qual a média é o ponto central ($A = \theta - x$ e $B = \theta + x$, com $x = \text{constante}$).

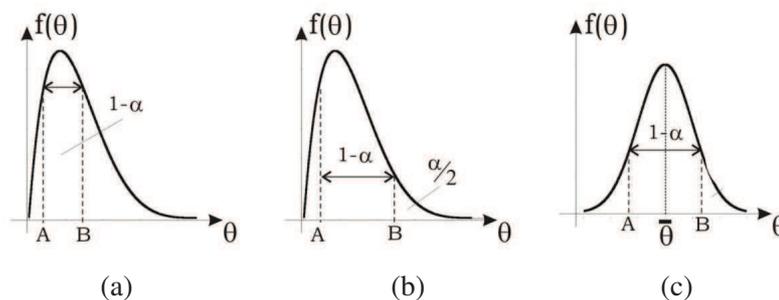


Figura 8 Intervalos de credibilidade
Fonte: Carvalho (2009)

Em geral, é desejado o intervalo de credibilidade com o menor comprimento possível e de mais alta densidade de probabilidade (HPD, do inglês highest probability density), apresentado na Figura 8a. Principalmente quando a distribuição de θ for assimétrica, o intervalo HPD se torna uma alternativa interessante, pois garante que o intervalo seja mínimo (CARVALHO, 2009). Segundo Bonat et

al. (2012), encontrar o intervalo HPD considerando uma distribuição assimétrica é um exemplo comum em inferência estatística.

Para obter o intervalo HPD para uma função de densidade, deve-se considerar que os limites (extremos) do intervalo tenham o mesmo valor de densidade ($f(A)=f(B)$ na Figura 8a). Assim, seleciona-se o menor de todos os intervalos possíveis para o nível desejado (BONAT et al., 2012). Em geral, o intervalo HPD minimiza a amplitude do intervalo a um dado nível de credibilidade, fornecendo estimativas mais confiáveis (CARVALHO, 2009).

A construção dos intervalos HPD, geralmente, é feita baseada em amostras de uma distribuição. Nesse caso, assume-se que a distribuição desejada é aproximada pela empírica da amostra. Em seguida, com base na propriedade de que o HPD é o menor de todos os intervalos, a implementação encontra os limites do menor intervalo com os valores da amostra ordenada, considerando o nível de credibilidade desejado (BONAT et al., 2012).

Para obter o intervalo HPD, pode-se utilizar a função implementada `HPDinterval()` do pacote `coda` do software R (R DEVELOPMENT CORE TEAM, 2013).

2.7.1 Uso de intervalos de credibilidade HPD na PCP

Com a implementação de um comitê de redes, é possível levar em consideração alguma incerteza sobre as predições de modelos de ANN. Para isso, são construídos intervalos de predição, a partir da distribuição de probabilidades empírica das predições das redes especialistas do comitê, para cada amostra do conjunto de teste. No entanto, nem sempre a distribuição empírica das predições é simétrica e, portanto, para a obtenção do intervalo de predição para cada amostra, deve-se considerar o intervalo de credibilidade de predições de menor intervalo

com máxima densidade de probabilidade, isto é, o intervalo HPD (BONAT et al., 2012).

Utilizando-se os intervalos HPD, é possível calcular a PCP e, assim, determinar qual a taxa de acerto de predição do modelo de comitê de redes. Os intervalos HPD são usados, então, como uma forma de quantificar a PCP. A construção do intervalo HPD, para cada amostra do conjunto de teste, é obtida da seguinte forma:

- uma distribuição empírica é construída para N predições de EMAN de um ingrediente específico, obtidas pelas N redes neurais do comitê;
- a partir da distribuição empírica, é considerada a amplitude que inclua os pontos de máxima densidade e tenha o mínimo comprimento, ou seja, o intervalo HPD. Na construção deste intervalo considera-se um nível $1 - \alpha$ de credibilidade.

3 SOFTWARES UTILIZADOS

- FANNTOOL

FANNTOOL é um software gratuito, desenvolvido com base em uma biblioteca de rede neural livre de código aberto em C. No entanto, essa biblioteca está disponível em outras linguagens de programação. Seu nome é uma abreviação de *Fast Artificial Neural Network*. Esse software permite a implementação de redes neurais artificiais multicamadas, totalmente e parcialmente conectadas. É um software versátil, bem documentado, rápido e fácil de usar, sem ser necessária a implementação de algoritmos de programação. No site <http://code.google.com/p/fanntool/> estão disponíveis maiores detalhes sobre este software.

- MATLAB

MATLAB é um ambiente de trabalho usado para implementar e analisar algoritmos numéricos. Seu nome é uma abreviação de *MATrix LABORatory* (MATLAB), sendo baseado no uso de matrizes e, desse modo, a maioria das operações de análise numérica podem ser realizadas usando álgebra linear. Foi desenvolvido por “The Mathworks, Inc” no final da década de 70, contendo constantes atualizações e aperfeiçoamentos. O conjunto de toolboxes, que podem ser organizados conforme necessidades específicas, tornam o MATLAB uma poderosa ferramenta de análise de dados. A Neural Network Toolbox do MATLAB possibilita a aplicação dos mais variados tipos de redes neurais além de permitir a modificação de redes já existentes. Mais informações encontram-se em <http://www.mathworks.com/>.

- Software R

O software R é gratuito e de código aberto que propicia excelente ambiente para análises estatísticas e produções de gráficos de alta qualidade. É um software livre e está disponível para diversos sistemas operacionais (Windows, UNIX/Linux, MacOS). O conjunto de pacotes do R é uma de suas principais características. A instalação padrão do R vem com diversos pacotes, sendo possível adicionar novos pacotes para finalidades específicas sempre que se fizer necessário. Maiores informações e diversos materiais sobre o software podem ser encontrados em <http://www.r-project.org>.

Neste trabalho, todas as análises estatísticas foram realizadas com o software R (R DEVELOPMENT CORE TEAM, 2013). Para a implementação das redes neurais, inicialmente, foi utilizado o software FANNTOOL (Artigo 1). Com o intuito de obter uma maior flexibilidade na implementação das redes foi utilizado o MATLAB, o qual oferece uma eficiente e flexível toolbox ANN (Artigo 2).

REFERÊNCIAS

- AHMADI, H. et al. Prediction model for true metabolizable energy of feather meal and poultry offal meal using group method of data handling-type neural network. **Poultry Science**, Champaign, v. 87, p. 1909-1912, Sept. 2008.
- AHMADI, H.; MOTTAGHITALAB, M.; NARIMAN-ZADEH, N. Group method of data handling-type neural network prediction of broiler performance based on dietary metabolizable energy, methionine, and lysine. **Journal of Applied Poultry Research**, Athens, v. 16, n. 4, p. 494-501, 2007.
- ALBINO, L. F. T. et al. Utilização de diferentes sistemas de avaliação energéticas dos alimentos na formulação de rações para frangos de corte. **Revista da Sociedade Brasileira de Zootecnia**, Viçosa, MG, v. 21, n. 6, p. 1037-1046, nov./dez. 1992.
- ALVARENGA, R. R. et al. Energetic values of feedstuffs for broilers determined with in vivo assays and prediction equations. **Animal Feed Science and Technology**, Amsterdam, v. 168, p. 257-266, Sept. 2011.
- BACAUSKIENE, M.; VERIKAS, A. Selecting salient features for classification based on neural network committees. **Pattern Recognition Letters**, Amsterdam, v. 25, p. 1879-1891, Dec. 2004.
- BALCEAN, S.; OOGHE, H. **Alternative methodologies in studies on business failure: do they produce better results than the classical statistical methods?** Gent: Universiteit Gent, 2004. 13 p.
- BARBOSA, A. H.; FREITAS, M. S. R.; NEVES, F. A. Confiabilidade estrutural utilizando o método de Monte Carlo e redes neurais. **REM: Revista Escola de Minas**, Ouro Preto, v. 58, p. 247-255, jul./set. 2005.
- BASHEER, I. A.; HAJMEE, M. Artificial neural networks: fundamentals, computing, design, and application. **Journal of Microbiological Methods**, Amsterdam, v. 43, p. 3-31, Dec. 2000.
- BISHOP, C. M. **Neural networks for pattern recognition**. Oxford: Clarendon, 1995. 482 p.

- BOLZAN, A. C.; MACHADO, R. A. F.; PIAIA, J. C. Z. Egg hatchability prediction by multiple linear regression and artificial neural networks. **Brazilian Journal of Poultry Science**, Campinas, v. 10, p. 97-102, Apr./June 2008.
- BONAT, W. H. et al. **Métodos computacionais para inferência estatística**. João Pessoa: Associação Brasileira de Estatística, 2012. 289 p.
- BORGES, F. M. O. et al. Equações de regressão para estimar valores energéticos de grão de trigo e seus subprodutos para frangos de corte, a partir de análises químicas. **Arquivo Brasileiro de Medicina Veterinária e Zootecnia**, Belo Horizonte, v. 55, p. 734-746, dez. 2003.
- BRAGA, A. P.; CARVALHO, A. C. P. L. F.; LUDERMIR, T. B. **Redes neurais artificiais: teoria e aplicações**. Rio de Janeiro: LTC, 2000. 250 p.
- BRITO JÚNIOR, C. A. R. et al. Redes neurais artificiais aplicadas para a predição do comportamento dinâmico-mecânico de compósitos de matriz epóxi reforçados com fibras de carbono. **Matéria**, Rio de Janeiro, v. 12, n. 2, p. 346-357, 2007.
- CAMPOS, J. R. et al. Implementação de redes neurais artificiais utilizando a linguagem de programação Java. In: BRAZILIAN CONFERENCE ON DYNAMICS, CONTROL AND THEIR APPLICATIONS - DINCON, 9., 2010, Serra Negra. **Anais** . . . Serra Negra: Dincon, 2010. 1 CD-ROM.
- CANUTO, A. M. P. et al. Investigating the influence of the choice of ensemble members in accuracy and diversity of selection-based and fusion-based methods for ensembles. **Pattern Recognition Letters**, Amsterdam, v. 28, n. 4, p. 472-486, Mar. 2007.
- CARVALHO, D. C. O. **Gráficos de controle bayesianos em alguns processos da família exponencial**. 2009. 60 p. Dissertação (Mestrado em Matemática e Estatística) - Universidade Federal do Pará, Belém, 2009.
- COCHRAN, W. G. The combination of estimates from different experiments. **Biometrics**, Washington, v. 10, p. 101-129, 1954.
- COSTA, E. J. X. Inteligência artificial aplicada à zootecnia. **Revista Brasileira de Zootecnia**, Viçosa, MG, v. 38, n. 4, p. 390-396, jul./ago. 2009.

CYBENKO, G. **Continuous valued neural network with two hidden layers are sufficient**: technical report. Medford: Tufts University, 1988.

FAGARD, R. H.; STAESSEN, J. A.; THIJS, L. Advantages and disadvantages of the meta-analysis approach. **Journal of Hypertension**, London, v. 14, n. 2, p. 9-13, Sept. 1996. Supplement.

FERNEDA, E. Redes neurais e sua aplicação em sistemas de recuperação de informação. **Ciência da Informação**, Brasília, v. 35, n. 1, p. 25-30, jan./abr. 2006.

FRANÇA, J. A. et al. Uma implementação do algoritmo levenberg-marquardt dividido para aplicações em visão computacional. **Semina: Ciências Exatas/Tecnológicas**, Londrina, v. 30, n. 1, p. 51-62, 2009.

GHEYAS, I. A.; SMITH, L. S. A novel neural network ensemble architecture for time series forecasting. **Neurocomputing**, New York, v. 74, p. 3855-3864, Nov. 2011.

HAYKIN, S. **Neural networks a comprehensive foundation**. 3rd ed. Upper Saddle River: Prentice Hall, 2007.

HAYKIN, S. **Neural networks and learning machines**. 3rd ed. Upper Saddle River: Prentice Hall, 2009. 936 p.

HAYKIN, S. **Redes neurais, princípios e prática**. 2. ed. Porto Alegre: Bookman, 2001. 900 p.

KENARI, S. A. J.; MASHOHOR, S. Robust committee machine for water saturation prediction. **Journal of Petroleum Science and Engineering**, Amsterdam, v. 104, p. 1-10, Apr. 2013.

KORDOS, M. et al. Neural network committees optimized with evolutionary methods for steel temperature control. **Lecture Notes in Artificial Intelligence**, Berlin, v. 6922, p. 42-51, 2011.

LIMA, C. A. M. **Comitê de máquinas**: uma abordagem unificada empregando máquinas de vetores-suporte. 2004. 342 p. Tese (Doutorado em Engenharia Elétrica) - Universidade Estadual de Campinas, Campinas, 2004.

LIPNICKAS, A. Adaptive committees of neural classifiers. **Information Technology and Control**, New York, v. 37, n. 3, p. 205-210, July/Sept. 2008.

LOVATTO, P. A. et al. Meta analysis in scientific research: a methodological approach. **Brazilian Journal of Animal Science**, Brasília, v. 36, p. 285-294, July 2007. Supplement.

MACLIN, R.; SHAVLIK, J. Combining the predictions of multiple classifiers: using competitive learning to initialize neural networks. In: INTERNATIONAL JOINT CONFERENCE ON ARTIFICIAL INTELLIGENCE, 14., 1995, Montreal. **Proceedings...** Montreal: M. Kaufmann, 1995. p. 524-530.

MARIANO, F. C. M. Q. et al. Equações de predição de valores energéticos de alimentos obtidas utilizando meta-análise e componentes principais. **Ciência Rural**, Santa Maria, v. 42, n. 9, p. 1634-1640, set. 2012.

MARIANO, F. C. M. Q. et al. Prediction of the energy values of feedstuffs for broilers using meta-analysis and neural networks. **Animal**, Cambridge, v. 7, n. 9, p. 1440-1445, Sept. 2013.

MCCULLOCH, W. S.; PITTS, W. A logical calculus of the ideas immanent in nervous activity. **Bulletin of Mathematical Biophysics**, New York, v. 5, p. 115-133, 1943.

NAGATA, A. K. et al. Energia metabolizável de alguns alimentos energéticos para frangos de corte, determinada por ensaios metabólicos e por equações de predição. **Ciência e Agrotecnologia**, Lavras, v. 28, n. 3, p. 668-677, maio/jun. 2004.

NASCIMENTO, G. A. J. et al. Equações de predição para estimar os valores energéticos de alimentos concentrados de origem vegetal para aves utilizando a metanálise. **Revista Brasileira de Zootecnia**, Viçosa, MG, v. 38, n. 7, p. 1265-1271, dez. 2009.

NASCIMENTO, G. A. J. et al. Equações de predição para estimar valores da energia metabolizável de alimentos concentrados energéticos para aves utilizando meta-análise. **Arquivo Brasileiro de Medicina Veterinária e Zootecnia**, Belo Horizonte, v. 63, n. 1, p. 222-230, fev. 2011.

NICHOLLS, J. G. et al. **From neuron to brain**. 5th ed. Sunderland: Sinauer Associates, 2012. 621 p.

NILSSON, N. J. **Learning machines: foundations of trainable pattern-classifying systems**. New York: MacGraw-Hill, 1965. 137 p.

OKUT, H. et al. Prediction of body mass index in mice using dense molecular markers and a regularized neural network. **Genetical Research**, London, v. 93, n. 3, p. 189-201, June 2011.

OLIVEIRA, R. C. et al. Aplicação de um comitê de redes neurais artificiais para a solução de problemas inversos em transferência radiativa. **TEMA - Tendências em Matemática Aplicada e Computacional**, São Carlos, v. 11, n. 2, p. 171-182, maio/ago. 2010.

PERAI, A. H. et al. A comparison of artificial neural networks with other statistical approaches for the prediction of true metabolizable energy of meat and bone meal. **Poultry Science**, Champaign, v. 89, p. 1562-1568, July 2010.

R DEVELOPMENT CORE TEAM. **R: a language and environment for statistical computing**. Vienna: R Foundation for Statistical Computing, 2013. Software.

RODRIGUES, P. B. et al. Valores energéticos da soja e subprodutos da soja, determinados com frangos de corte e galos adultos. **Revista Brasileira de Zootecnia**, Viçosa, MG, v. 31, n. 4, p. 1771-1782, jul./ago. 2002.

RODRIGUES, P. B. et al. Valores energéticos do milho, do milho e subprodutos do milho, determinados com frangos de corte e galos adultos. **Revista Brasileira de Zootecnia**, Viçosa, MG, v. 30, n. 6, p. 1767-1778, nov./dez. 2001.

RUMELHART, D. E.; MCCLELLAND, J. L. **Parallel distributed processing: explorations in the microstructures of cognition**. Cambridge: MIT, 1986. v. 1, 567 p.

RUSTEMPASIC, I.; CAN, M. Diagnosis of Parkinson's disease using principal component analysis and boosting committee machines. **Southeast Europe Journal of Soft Computing**, Ilidza, v. 2, n. 1, p. 102-109, Mar. 2013.

SAUVANT, D. et al. Meta-analyses of experimental data in animal nutrition. **Animal**, Cambridge, v. 2, n. 8, p. 1203-1214, Aug. 2008.

SEGATTO, Ê. C. et al. Alto desempenho na proteção diferencial de transformadores de potência com a utilização de redes neurais artificiais. **Controle & Automação**, Campinas, v. 14, n. 3, p. 309-320, jul./set. 2003.

SIWEK, K.; OSOWSKI, S.; SZUPILUK, R. Ensemble neural network approach for accurate load forecasting in a power system. **International Journal of Applied Mathematics and Computer Science**, Zielona Góra, v. 19, p. 303-315, June 2009.

THOMÉ, A. C. G. **Redes neurais**: uma ferramenta para KDD e data mining. Disponível em: <http://equipe.nce.ufrj.br/thome/grad/nn/mat_didatico/apostila_kdd_mbi.pdf>. Acesso em: 20 mar. 2012.

TSYMBAL, A.; PECHENIZKIY, M.; CUNNINGHAM, P. Diversity in search strategies for ensemble feature selection. **Information Fusion**, New York, v. 6, p. 83-98, Mar. 2005.

WAN, H. F. et al. Prediction of true metabolizable energy from chemical composition of wheat milling by-products for ducks. **Poultry Science**, Champaign, v. 88, p. 92-97, Jan. 2009.

YEE, D.; PRIOR, M. G.; FLORENCE, L. Z. Development of predictive models of laboratory animal growth using artificial neural networks. **Computer Applications in the Biosciences**, Liverpool, v. 9, p. 517-522, Oct. 1993.

ZHAO, F. et al. Predicting metabolizable energy of normal corn from its chemical composition in adult pekin ducks. **Poultry Science**, Champaign, v. 87, p. 1603-1608, Aug. 2008.

ZONTA, M. C. M. et al. Energia metabolizável de ingredientes protéicos determinada pelo método de coleta total e por equações de predição. **Ciência e Agrotecnologia**, Lavras, v. 28, n. 6, p. 1400-1407, nov./dez. 2004.

SEGUNDA PARTE
Artigos

ARTIGO 1: Prediction of the energy values of feedstuffs for broilers using meta-analysis and neural networks

Redigido conforme as normas da revista *Animal* (Cambridge) - versão final.

F.C.M.Q. Mariano¹, C.A. Paixão², R.R. Lima¹, R.R. Alvarenga³, P.B. Rodrigues³ and G.A.J. Nascimento⁴

¹Department of Exact Science, Federal University of Lavras, PO Box 3037, 37200-000 Lavras, Minas Gerais, Brazil; ²Applied Mathematics School, Getulio Vargas Foundation, 22250-900 Rio de Janeiro, Rio de Janeiro, Brazil; ³Department of Animal Science, Federal University of Lavras, PO Box 3037, 37200-000 Lavras, Minas Gerais, Brazil; ⁴Department of Animal Science, Federal University of Ceará, PO Box 12168, 60455-970 Fortaleza, Ceará, Brazil.

Corresponding author: Flávia Cristina M. Q. Mariano. E-mail: flaviaqz@gmail.com

(Received 8 June 2012; Accepted 11 March 2013)

Several researchers have developed prediction equations to estimate the metabolisable energy (ME) of energetic and protein concentrate feedstuffs used in diets for broilers. The ME is estimated by considering CP, ether extract, ash and fibre contents. However, the results obtained using traditional regression analysis methods have been inconsistent and new techniques can be used to obtain better estimate of the feedstuffs' energy value. The objective of this paper was to implement a multi-layer perceptron network to estimate the nitrogen-corrected metabolisable energy (AMEn) values of the energetic and protein concentrate feeds, generally used by the poultry feed industry. The concentrate feeds were from plant origin. The dataset contains 568 experimental results, all from Brazil. This dataset was separated into two parts: one part with 454 data, which was used to train, and the other one with 114 data, which was used to evaluate the accuracy of each implemented network. The accuracy of the models was evaluated on the basis of their values of mean squared error, R^2 , mean absolute deviation, mean absolute percentage error

and bias. The 7-5-3-1 model presented the highest accuracy of prediction. It was developed an Excel[®] AMEn calculator by using the best model, which provides a rapid and efficient way to predict the AMEn values of concentrate feedstuffs for broilers.

Keywords: avian production; broilers; metabolisable energy; multilayer perceptron.

Implications

It is difficult and expensive to estimate the nitrogen-corrected metabolisable energy (AMEn) values of the feedstuffs used for broilers. The results of this study demonstrate that the implementation of multilayer perceptron networks (MLP) in a meta-analysis is suitable to estimate these energy values. Furthermore, a calculator was created on the basis of the results of the MLP, which allows an efficient way of predicting the AMEn values.

Introduction

Knowledge of the chemical composition and metabolisable energy (ME) of feedstuffs is necessary to provide an adequate supply of nutrients and energy for animals. A variety of feedstuffs and their by-products are used in diets, and it is important to know accurately the dietary nutrients that each contains. The energy content of feedstuffs may be determined using metabolic bioassays (Rodrigues *et al.*, 2001; Zhao *et al.*, 2008; Wan *et al.*, 2009), which are onerous and time-consuming. Alternative ways to obtain these values include using the composition of feedstuffs and nutritional composition tables, and prediction equations based on the chemical composition of the feedstuffs.

Several studies have developed prediction equations to estimate the ME using regression methods. However, the results obtained using traditional regression methods have been inconsistent (Alvarenga *et al.*, 2011). An interesting way to obtain prediction equations that yield more consistent results is to combine information derived from data collected under different but related conditions. This method uses regression theory by considering the meta-analysis principle.

Meta-analysis is a relevant method for summarising and quantifying knowledge acquired through previously published research (Sauvant *et al.*, 2008). Although the papers involve the same subject, in meta-analyses it is important that homogeneous groups of papers are formed. This is one of the greatest difficulties in the development of meta-analysis (Fagard *et al.*, 1996; Lovatto *et al.*, 2007). In some studies, homogeneous groups of papers used to develop prediction equations for the AMEn of broiler feedstuffs were obtained by combining levels of different factors (Nascimento *et al.*, 2009 and 2011) or by using the multivariate technique of principal components (Mariano *et al.*, 2012).

Neural Networks is a term that denotes sets of connectionist models inspired by the neurological structures and processing function of the central nervous system of living beings, being an adaptive system that changes its structure based on external or internal information that flows through the network (Bishop, 1995; Haykin, 2007). In an artificial neural network (ANN), a neuron processes the weighted inputs and computes a single output by using an activation function. As an ANN consists of an interconnected group of artificial neurons, a neuron processes the received information from other neurons, which are called inputs. Although the neuron is a computational unit, neurons may be combined into layers to create an efficient network that can learn to distinguish behavioural patterns (Haider and Hanif, 2009).

The ANN has some advantages. The ANN does not need restrictive assumptions and it has the ability to learn general solutions on the basis of the data given for training (adaptive learning), to create its own organisation or representation of the information received during learning time and to analyse complex patterns. Although ANN is considered robust and flexible, there are some drawbacks in its use: it requires a large and high-quality training dataset; it is necessary to compare different architectures to select the best; the variables used in an ANN must be carefully selected a priori; and there is risk of overfitting (Balcean and Ooghe, 2004).

Artificial neural network can be applied with different objectives, such as pattern recognition systems, data processing, function approximation and clustering. The ANN has also been used as a form of prediction. Gheyas and Smith (2011) proposed an ANN for time series forecasting. Okut *et al.* (2011) predicted body mass index using a regularised neural network. ANN has also been used to predict ME (Ahmadi *et al.*, 2007 and 2008; Perai *et al.*, 2010). The multilayer perceptron neural network (MLP) has been successfully applied to predict the true metabolisable energy (TMEn) values of meat and bone meal samples (Perai *et al.*, 2010).

An MLP consists of a set of source nodes, which form the input layer of the network. All others layers are composed of neurons that present computational capacity, as shown in Figure 1. This type of neural network is a progressive network in which the outputs of the neurons are connected only to the inputs of neurons of the next layer, without connections within the layers. Consequently, the input signal propagates through the network, layer by layer, in a progressive direction (Bishop, 1995; Haykin, 2007).

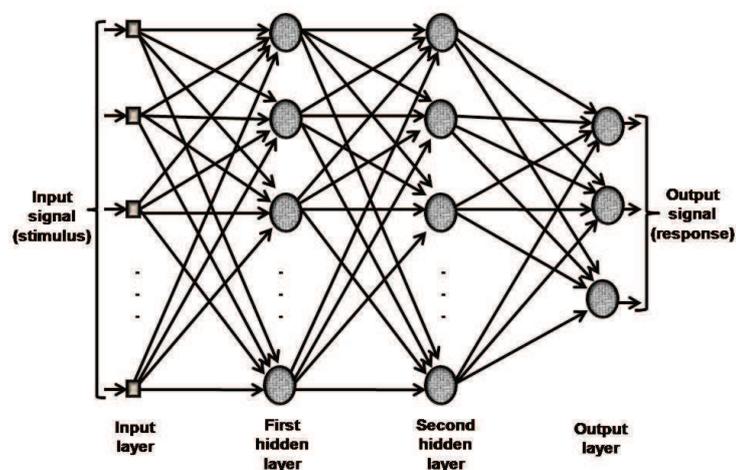


Figure 1 Structure of a multilayer perceptron with two hidden layers.

The objective of this study was to estimate the AMEn values of the energetic and protein concentrate feeds of plant origin used for broilers by using an MLP in a metaanalysis study. It will be created an Excel[®] AMEn calculator, which can be used by the animal nutritionists to predict AMEn for feedstuff samples.

Material and methods

Data source

The database used for the meta-analysis has been catalogued and described previously by Nascimento *et al.* (2009 and 2011). A bibliographical review of studies carried out in Brazil was performed to collect information concerning the AMEn values and chemical composition of feedstuffs for broilers. All feedstuffs were of plant origin. The database included a total of 568 experimental results corresponding to 370 energetic concentrate and 198 protein concentrate samples evaluated in experiments executed from 1967 to 2007 (Table 1).

Table 1 *Number of experimental results considered in this study grouped by concentrate and feedstuff types*

Concentrate ^a	Feedstuffs	Number of experimental results	Total
Energetic	Maize	168	370
	Maize by-product	29	
	Other	173	
Protein	Soya bean meal	59	198
	Other soya bean by-product	59	
	Other	80	

^a Concentrate of plant origin commonly used in poultry diet.

The AMEn, the chemical composition values (CP; ether extract without acid hydrolysis - EE; ash; crude fibre - CF), the classification of the feedstuff category (1 - energetic concentrate; 2 - protein concentrate), the specification of the ingredient (1 - maize or soya bean meal; 2 - maize byproduct or other soya bean by-product; 3 - other feedstuff) and the type of animal used in the bioassay (1 - chick; 2 - rooster) were defined for each feeds. The AMEn was determined on the basis of a bioassay. The NDF and ADF values were not considered because this information was not available in the vast majority of experimental results. The list of references used for the meta-analysis is provided in the Supplementary Material; references for the energetic concentrates are provided in Supplementary Material 1; and references for the protein concentrates are provided in Supplementary Material 2.

Model development

The AMEn values of concentrate feedstuffs were predicted from an MLP using the meta-analysis principle. The dataset, part of which is presented in Table 2, was randomly partitioned into two separate subsets: one, comprising 80% (454 studies) of the data, was used as training set in the development of the network; the second one, comprising 20% (114 studies) of the data, was used as test set. The training data should not be used to test the network because the network can

Table 2 *Data sample lines used to develop the multilayer perceptron network model for the AMEn values (on dry matter basis) of feedstuffs for broilers*

n	Inputs							Output
	CP ^(a)	EE ^(a)	Ash ^(a)	CF ^(a)	CAT	ING	ANIMAL	AMEn (kcal/Kg)
1	10.92	4.03	1.81	1.92	1	1	1	3573
2	10.75	3.53	1.82	2.00	1	1	1	3560

302	11.80	2.40	3.20	6.50	1	2	1	2700
303	1.76	0.06	3.01	17.67	1	3	2	2340

567	32.02	10.4	4.95	2.50	2	3	2	1620
568	32.02	10.4	4.95	2.50	2	3	2	2310

AMEn=nitrogen-corrected metabolisable energy; EE=ether extract; CF=crude fibre; n=number of data; CAT=classification of the feedstuff category:

1 - energetic concentrate; 2 - protein concentrate; ING=specification of the ingredient: 1 - maize or soybean meal; 2 - maize by-product or other soybean by-product; 3 - other feedstuff; ANIMAL=type of animal used in the bioassay: 1 - chick; 2 - rooster.

^(a) Values expressed on dry matter basis.

memorise the data pattern and the test procedure would not be reliable.

The input parameters of the implemented MLP were feedstuff categories (CAT), concentrate ingredients (ING), type of animals used in the bioassay (ANIMAL), CP, EE, MM and CF. The AMEn column (Table 2) was the values of desired output.

Different structures were implemented with the objective of selecting the best one. In the general way, the structured was defined as 7-p-q-1, which represents: seven inputs, p neurons in the first hidden layer, q neurons in the second hidden layer and one output. Values from 1 to 10 were considered for both p and q. The MLP proposed contains two hidden layers because this structure allows any function approximation (Cybenko, 1988).

A backpropagation algorithm was used to train the network. This algorithm is based on the error-correction learning rule, which is propagated backward

from the output layer to hidden layers of the ANN (Rumelhart *et al.*, 1986). Basically, this algorithm iteratively adjusts the weights to the network randomly, presenting examples to the neural network as an input signal (Albuquerque *et al.*, 2009). The algorithm minimises the mean square error (m.s.e.) of prediction, whose error is the difference between the desired outcome from the data input and the predicted response from the output neuron.

The software FANN TOOL 1.2 (<http://code.google.com/p/fanntool/>) was used to implement the networks. The selected training algorithm was the iRPROP, which was described by Igel and Husken (2000). The initial weights were randomly defined by the software, and the network was trained up to 500,000 epochs. The activation function for the hidden neurons was the fast “sigmoid-like” function and for the output neuron was the periodical cosine function. The learning rate and momentum for network training were set, respectively, at 0.7 and 0. The goodness of fit of the model and the accuracy of the predicted AMEn were evaluated using the training and testing data. The measures used in this process were as follows: Coefficient of determination (R^2), m.s.e., mean absolute deviation (MAD), mean absolute percentage error (MAPE) and bias, as defined in Bolzan *et al.* (2008) and Perai *et al.* (2010).

Results and discussion

Two MLP models, which presented the best accuracy in the prediction of the AMEn values, were selected. A summary of statistical results associated with these two MLP models is shown in Table 3. These statistics indicate forecasting error measurements based on the difference between observed and predicted values. The closer to one the R^2 value, the closer to zero the bias value, and the lower m.s.e., MAPE and MADE values, the more accurate is the model.

Table 3 Statistics used to check the goodness-of-fit and the accuracy of the two best multilayer perceptron network models

Statistics	7-5-5-1 model		7-5-3-1 model	
	MLP training	MLP testing	MLP training	MLP testing
R ² (%)	0.89	0.83	0.91	0.86
m.s.e.	58,531.42	104,091.40	44,933.39	86,725.77
MAD	160.12	215.58	152.00	210.27
MAPE (%)	5.76	8.39	5.21	8.30
Bias	1.00	21.54	-0.31	-26.34

R² - coefficient of determination; MAD - mean absolute deviation; MAPE - mean absolute percentage error; m.s.e. - mean square error.

The 7-5-3-1 model had lower values of m.s.e., MAD and MAPE than the 7-5-5-1 model, both training and test dataset. Furthermore, the first model presented values of R² closer to one, and the MLP training bias closer to zero, than the second one. Thus, the 7-5-3-1 model was considered more suitable to predict the AMEn for both the energetic and protein feedstuffs for broilers.

The selected MLP model has higher prediction accuracy (higher R² values) than the models obtained by Nascimento *et al.* (2009) and Mariano *et al.* (2012). These authors had obtained equations to predict the AMEn for both the energetic and protein feedstuffs for broilers, by using meta-analysis and regression methods. Their models presented R² values approximately equal to 0.83 and 0.74, respectively.

In general, there was better prediction of AMEn using the training dataset than testing dataset (Table 3). Ahmadi *et al.* (2007) had also found these same differences between the two datasets. The values for the m.s.e. are into the interval presented by Perai *et al.* (2010). These authors found m.s.e. values between 2,338.1875 and 91,335.17. Perai *et al.* (2010) obtained m.s.e. values for the ANN model lower than the m.s.e. values presented in the Table 3. However, they used a specific protein feedstuff. In this paper, 568 AMEn values of energetic and protein feedstuffs were used. These feedstuffs have a higher variability in the

AMEn values, and this leads to an increase in the m.s.e.

It is important to note that a larger dataset was used in the ANN procedure because this improves the chance of obtaining better adjustments (Bishop, 1995). This is important because the statistics used to check for ANN convergence are usually better estimated when the dataset is larger. Ahmadi *et al.* (2008) used just 30 raw data lines consisting of 12 feather meal (FM) and 18 poultry offal meal (POM) samples to train a group method of data handling-type neural network (GMDH-type NN). Another group with seven data lines (three FM and four POM) were used as validation set of the ANN. Perai *et al.* (2010) used 34 raw lines and only one feedstuff (meat and bone meal-MBM). Although Ahmadi *et al.* (2008) and Perai *et al.* (2010) used a small dataset, their prediction was accurate. This most likely occurred because they used specific feedstuffs. In this study, an ANN was developed to predict the AMEn for energetic and protein concentrates, though it is important to emphasise that the implemented MLP is more suitable to predict the AMEn of the feedstuffs commonly used in poultry feeds, such as maize, sorghum, soya bean, soya bean meal, wheat and wheat meal.

The predicted *v.* observed values of the test data are presented in Figure 2. Points closer to the identity line indicate that the predicted values are similar to the observed data. Therefore, the predicted values obtained using MLP tend to follow the same patterns as the observed data. These results indicate that the AMEn values were accurately predicted.

It can be observed in the Figure 2 that there are few points that the predictions are quite off the mark. Some points indicate that the difference between the observed and the predicted AMEn values (error) is equal to plus or minus 1000 kcal. These points refer to the feedstuffs that are not commonly used in poultry feeds. Moreover, there is a small number of these feedstuffs in the dataset,

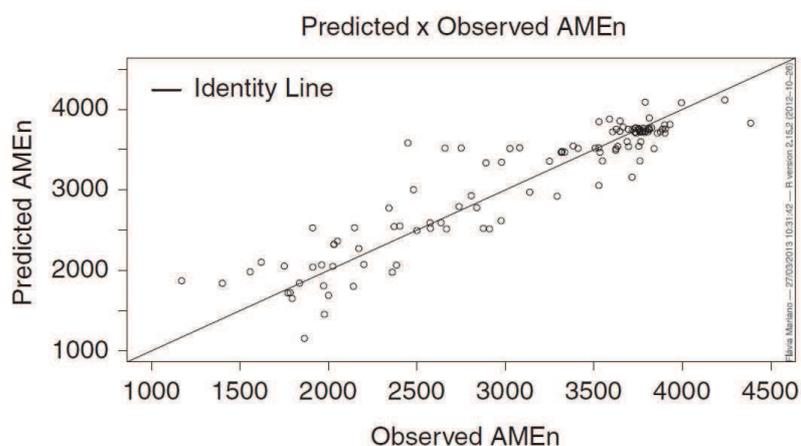


Figure 2 Predicted *v.* observed AMEn values obtained from the test dataset for concentrate feedstuffs for broilers.

which affect the accuracy of prediction. For these feedstuffs, there is no processing pattern, resulting in a considerable variation in their chemical composition and, consequently, in their energy value (Rodrigues *et al.*, 2001; Moreira *et al.*, 2002; Brunelli *et al.*, 2006).

No general criterion exists to define the number of neurons in the hidden layer. In general, neural networks with few hidden neurons are preferred because they tend to have better generalisation power, thereby reducing the problem of overfitting. However, networks with few hidden neurons may not be able to model and learn the data in complex problems, and this can result in underfitting, that is, the network does not converge during training (Pereira, 1999; Calôba *et al.*, 2002).

In some of the cases where the number of neurons in the single hidden layer becomes high, the use of two or three layers may sometimes allow the number of neurons in the hidden layer to be reduced (Santos *et al.*, 2005). According to Wijayasekara *et al.* (2011), the chance of over-training a network increases with the number of neurons and the number of training epochs. Thus, as the number

of neurons increases, the possibility of the network describing the training data pattern exactly also increases. This is an undesirable ANN behaviour because the objective is to obtain an MLP that accurately predicts the AMEn values.

The prediction of energy values by using regression models considers just chemical composition variables (Zhao *et al.*, 2008; Wan *et al.*, 2009; Mariano *et al.*, 2012). The MLP can be fit by using other factors in addition to chemical composition, which is a great advantage. These factors can influence the AMEn and can improve its prediction. The type of the feedstuffs (energetic or protein concentrate), the specification of the ingredient (maize, soya bean meal and their similar products or others feedstuffs) and the type of animal used in the bioassay (chick or rooster) were considered to fit the MLP model, as these variables affect the AMEn variability.

The results obtained in this study support the findings of previous studies, in that the use of ANN has demonstrated promising results regarding predictions in the science poultry field, including the evaluation of broiler diets. Perai *et al.* (2010) compared the performance of three-layer feedforward ANN, partial least squares (PLS) and multiple linear regression (MLR) methods to predict the TMEn values of meat and bone meat samples based on their chemical composition. The results demonstrated that the ANN model outperformed the PLS and MLR models. A GMDH-type NN accurately predicted broiler performance on the basis of dietary metabolisable energy, methionine and lysine (Ahmadi *et al.*, 2007) and predicted the TMEn values of feather and POMs on the basis of their chemical composition (Ahmadi *et al.*, 2008).

Once selected the best MLP model, an Excel[®] AMEn Calculator was created (Figure 3). This AMEn Calculator enables the animal nutritionists to use this tool to predict AMEn for feedstuff samples. It is provided in Supplementary Ma-

Excel AMEn Calculator	
Prediction of AMEn depending on the chemical composition and specifications of the feedstuffs used by the poultry feed industry using neural network model.	
Nitrogen-corrected Metabolisable Energy - AMEn ^(a)	3523,33
AMEn value expressed on natural matter basis	3170,99
Dry Matter (%)	90
Crude Protein (%) ^{(a),(b)}	4,26
Ether Extract (%) ^{(a),(b)}	3,63
Ash (%) ^{(a),(b)}	2,48
Crude Fibre (%) ^{(a),(b)}	0,31
Category ^(c)	1
Ingredient ^(c)	3
Animal ^(c)	1
^(a) Values expressed on dry matter basis	
^(b) Ranges of the chemical composition values considered in the development of the neural network model:	
- PB (%): [1,47; 71,44]	- EE (%): [0,03; 26,21]
- Ash (%): [0,30; 12,61]	- FB (%): [0,02; 27,63]
^(c) Values of Category, Ingredient and Animal:	
Category:	
- Concentrate Energetic	= 1
- Concentrate Protein	= 2
Ingredient:	
- Maize or Soybean Meal	= 1
- Maize By-product or Other Soybean By-product	= 2
- Other feedstuff	= 3
Animal:	
- Chick	= 1
- Rooster	= 2

Figure 3 Excel[®] calculator to predict the AMEn values of concentrate feedstuffs for broilers.

terial 3. Furthermore, it is intended to leave it available in conjunction with the nutritional composition tables. Although only Brazilian data (preliminary study) were used in the development of the MLP, this calculator could be used with data from international origin.

In future studies, the dataset including other experimental results from Brazilian and other countries studies will be updated. The objective will be to improve the accuracy of the AMEn prediction for broilers by using the ANN model. Other types of ANN and additional techniques will be explored in order to optimise generalisation of the networks (e.g. cross-validation, early stopping and identification of important input parameters in building a MLP).

Conclusion

The MLP with the best performance has a structure of seven inputs, five neurons in the first hidden layer, three neurons in the second hidden layer and one neuron as output. This result revealed that use of MLP method is a promising approach for the accurate prediction of AMEn values of energetic and protein concentrate feeds used in broiler diets. The Excel[®] AMEn calculator developed in this study is an efficient and easy way to predict the AMEn values by using ANN.

Acknowledgements

The authors gratefully acknowledge financial supports provided by the Coordination Improvement of Higher Education Students (CAPES) and by the National Council for Scientific and Technological Development - Brazil (CNPq). The authors would also like to thank the editor Jaap van Milgen and the anonymous reviewers for their valuable comments and suggestions to improve the quality of this paper.

Supplementary materials

For supplementary materials referred to in this article, please visit <http://dx.doi.org/10.1017/S1751731113000712>.

References

Ahmadi H, MottaghitalabM and Nariman-Zadeh N 2007. Group method of data handling-type neural network prediction of broiler performance based on dietary metabolizable energy, methionine, and lysine. *Journal of Applied Poultry Research* 16, 494-501.

Ahmadi H, Golian A, Mottaghitlab M and Nariman-Zadeh N 2008. Prediction model for true metabolizable energy of feather meal and poultry offal meal using group method of data handling-type neural network. *Poultry Science* 87, 1909-1912.

Albuquerque VHC, Alexandria AR, Cortez PC and Tavares JMRS 2009. Evaluation of multilayer perceptron and self-organizing map neural network topologies applied on microstructure segmentation from metallographic images. *NDT & E International* 42, 644-651.

Alvarenga RR, Rodrigues PB, Zangeronimo MG, Freitas RTF, Lima RR, Bertechini AG and Fassani EJ 2011. Energetic values of feedstuffs for broilers determined with in vivo assays and prediction equations. *Animal Feed Science and Technology* 168, 257-266.

Balcean S and Ooghe H 2004. Alternative methodologies in studies on business failure: do they produce better results than the classical statistical methods? Working Paper of Faculty of Economics and Business Administration, Ghent University, Belgium, 40pp.

Bishop CM 1995. *Neural networks for pattern recognition*. Oxford University Press, Oxford, UK, 482pp.

Bolzan AC, Machado RAF and Piaia JCZ 2008. Egg hatchability prediction by multiple linear regression and artificial neural networks. *Brazilian Journal of Poultry Science* 10, 97-102.

Brunelli SR, Pinheiro JW, Silva CA, Fonseca NAN, Oliveira DD, Cunha GE and Souza LFA 2006. Feeding increasing defatted corn germ meal levels to broiler chickens. *Brazilian Journal of Animal Science* 35, 1349-1358.

Calôba GM, Calôba LP and Saliby E 2002. Cooperação entre redes neurais artificiais e técnicas clássicas para previsão de demanda de uma série de vendas de cerveja na Austrália. *Pesquisa Operacional* 22, 345-358.

Cybenko G 1988. Continuous valued neural network with two hidden layers are sufficient. Technical Report, Department of Computer Science, Tufts University, Medford, MA, USA.

Fagard RH, Staessen JA and Thijs L 1996. Advantages and disadvantages of the meta-analysis approach. *Journal of Hypertension* 14 (suppl. 2), 9-13.

Gheyas IA and Smith LS 2011. A novel neural network ensemble architecture for time series forecasting. *Neurocomputing* 74, 3855-3864.

Haider A and Hanif MN 2009. Inflation forecasting in Pakistan using artificial neural networks. *Pakistan economic and social review* 47, 123-138.

Haykin S 2007. *Neural networks - a comprehensive foundation*, 3rd edition. Prentice-Hall Inc., Upper Saddle River, NJ, USA.

Igel C and Hüsken M 2000. Improving the RPROP learning algorithm. *Proceedings of Second International Symposium on Neural Computing NC 2000*, 23-26 May, Berlin, Germany, pp. 115-21.

Lovatto PA, Lehen CR, Andretta I, Carvalho AD and Hauschild L 2007. Meta analysis in scientific research: a methodological approach. *Brazilian Journal of Animal Science* 36 (suppl.), 285-294.

Mariano FCMQ, Lima RR, Rodrigues PB, Alvarenga RR and Nascimento GAJ 2012. Prediction equations of energetic values of feedstuffs obtained using meta-analysis and principal components. *Ciência Rural* 42, 1634-1640.

Moreira I, Ribeiro CR, Furlan AC, Scapinello C and Kutschenko M 2002. Utilization of defatted corn germ meal on growing-finishing pigs feeding - digestibility and performance. *Brazilian Journal of Animal Science* 31, 2238-2246.

Nascimento GAJ, Rodrigues PB, Freitas RTF, Allaman IB, Lima RR and Reis Neto RV 2011. Prediction equations to estimate the AMEn values of protein feedstuffs for poultry utilizing meta-analysis. *Brazilian Journal of Animal Science* 40, 2172-2177.

Nascimento GAJ, Rodrigues PB, Freitas RTF, Bertechini AG, Lima RR and Pucci LEA 2009. Prediction equations to estimate the energy values of plant origin concentrate feeds for poultry utilizing the meta-analysis. *Brazilian Journal of Animal Science* 38, 1265-1271.

Okut H, Gianola D, Rosa GJM and Weigel KA 2011. Prediction of body mass index in mice using dense molecular markers and a regularized neural network. *Genetical Research* 93, 189-201.

Perai AH, Moghaddam HN, Asadpour S, Bahrapour J and Mansoori GH 2010. A comparison of artificial neural networks with other statistical approaches for the prediction of true metabolizable energy of meat and bone meal. *Poultry Science* 89, 1562-1568.

Pereira BB 1999. Introduction to neural networks in statistics. Center of Multivariate Analysis, Technical Report, Pennsylvania State University, Pennsylvania, USA.

Rodrigues PB, Rostagno HS, Albino LFT, Gomes PC, Barboza WA and Santana RT 2001. Energy values of millet, corn and corn byproducts, determined with broilers and adult cockerels. *Brazilian Journal of Animal Science* 30, 1767-1777.

Rumelhart DE, Hinton GE and Williams RJ 1986. Learning internal representations by error propagation. In *Paralleled distributed processing: explorations in the microstructure of cognition, vol. 1: foundations* (ed. DE Rumelhart and JL McClelland), pp. 318-362. The MIT Press, Cambridge, MA.

Santos AM, Seixas JM, Pereira BB and Medronho RA 2005. Using artificial neural networks and logistic regression in the prediction of Hepatitis A. *Revista Brasileira de Epidemiologia* 8, 117-126.

Sauvant D, Schmidely P, Daudin JJ and St-Pierre NR 2008. Meta-analyses of experimental data in animal nutrition. *Animal* 2, 1203-1214.

Wan HF, Chen W, Qi ZL, Peng P and Peng J 2009. Prediction of true metabolizable energy from chemical composition of wheat milling by-products for ducks. *Poultry Science* 88, 92-97.

Wijayasekara D, Manic M, Sabharwall P and Utgikar V 2011. Optimal artificial neural network architecture selection for performance prediction of compact heat exchanger with the EBaLM-OTR technique. *Nuclear Engineering and Design* 241, 2549-2557.

Zhao F, Zhang HF, Hou SS and Zhang ZY 2008. Predicting metabolizable energy of normal corn from its chemical composition in adult pekin ducks. *Poultry Science* 87, 1603-1608.

ARTIGO 2: Neural network committee to predict the AMEn of poultry feedstuffs

Redigido conforme as normas da revista Neural Computing and Applications - versão preliminar.

F.C.M.Q. Mariano¹⁴, R.R. Lima¹, R.R. Alvarenga², P.B. Rodrigues² and W.S. Lacerda³

¹Department of Exact Science, Federal University of Lavras, PO Box 3037, 37200-000 Lavras, Minas Gerais, Brazil; ²Department of Animal Science, Federal University of Lavras, PO Box 3037, 37200-000 Lavras, Minas Gerais, Brazil; ³Department of Computer Science, Federal University of Lavras, PO Box 3037, 37200-000 Lavras, Minas Gerais, Brazil.

⁴Corresponding author - E-mail: flaviaqz@gmail.com

Abstract

A committee of neural networks is the aggregation of two or more neural networks for making overall predictions that are supposedly more accurate than those obtained by the individual networks. The objective of this paper was to assign some uncertainty over the predictions of neural networks, using a network committee to estimate the nitrogen-corrected metabolizable energy (AMEn) values of the energetic and protein concentrate feedstuffs for broilers. The dataset used to implement each expert network contains 568 experimental results. Another dataset with 48 bioassay results was used as test data. From several implemented multilayer perceptrons, the networks that presented the best generalization performance were selected to constitute the committee. The percentage of correct predictions was used as the criterion to compare committees that contained different numbers of networks. The highest probability density (HPD) intervals were obtained for each

feedstuff in the test data in this comparison. The estimator that ensured more accurate predictions was selected. The accuracy was evaluated using the mean squared error, R^2 , the mean absolute deviation, the mean absolute percentage error and the Bias. The highest accuracy for predicting the AMEn values of concentrate feedstuffs for broilers was achieved by a committee with 1000 networks with the use of the mode of the empirical distribution obtained from 1000 estimated values of the AMEn.

Keywords: committee machine, ensemble estimators, HPD intervals, metabolizable energy.

1 Introduction

Artificial neural networks (ANNs) are part of a subfield of artificial intelligence or machine learning. The objectives of an ANN are to automatically acquire knowledge from examples and to improve the performance of a given task [10]. Developments in this area have occurred through the creation of several methods and computational techniques specifically inspired by human neurological structures and by processing functions of the central nervous systems of living beings. This comparison is useful because an ANN consists of an interconnected group of artificial neurons arranged in layers and connected by interconnections known as synaptic weights (synapses).

An ANN is an adaptive system that changes its structure based on external or internal information that flows through the network [4, 13]. The information received in each neuron is known as the input. Thus, the neuron is a computational unit that processes weighted inputs and computes a single output using an activation function. The outputs of the hidden layers are the inputs for other neurons. Although the neuron is a computational unit, the use of an appropriate learning

algorithm in the ANN can create an efficient network that can learn to distinguish behavioral patterns [12]. The learning algorithm is responsible for adjusting the synaptic weights, in which all of the acquired knowledge is stored [9, 13].

An ANN has the ability to learn through examples, to generalize learned information and to subsequently give coherent responses to new data. These abilities demonstrate that the application of an ANN goes beyond simply mapping relationships between input and output data. Through the use of examples, ANNs can extract information that is not given explicitly [5]. As a consequence, ANN models are often used to replace complex mathematical systems described by multiple differential equations [14]. ANNs have also been used for prediction [1-2, 21, 25].

A multilayer perceptron neural network (MLP) is a type of supervised learning ANN that has been successfully applied in areas with a variety of different objectives, such as pattern recognition systems, data processing, function approximation and clustering. An MLP consists of a set of source nodes that form the input layer of the network. All other layers are composed of neurons that present computational capacity. This type of neural network is a progressive network in which the neuron outputs are connected only to the inputs of neurons of the next layer, without other connections within the layers (Fig. 1). Consequently, the input signal propagates through the network, layer by layer, in a progressive direction [4, 13].

The combination of several networks constitutes a committee machine, or more precisely, a committee of networks (CN). The CN methodology is based on the principle of divide and conquer [13] and can function as a supervised or unsupervised learning method. The use of a CN seeks to enhance the generalization of estimators, avoiding data overfitting and network convergence to a local minimum.

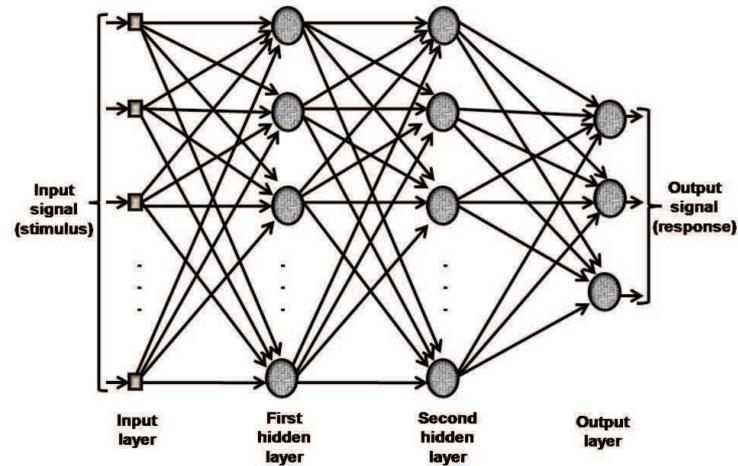


Fig. 1 Structure of a multilayer perceptron with two hidden layers

As its name indicates, a CN provides a fusion of knowledge acquired by more than one component to arrive at an overall solution, which is supposedly superior to that achieved by any one component acting alone [13, 17, 28].

Every individual CN component is an ANN with different ability to find a generalized solution to the problem; each component is known as an expert network, or simply an expert. As Lima [17] asserted, each committee component must exhibit good performance when used alone and must not be correlated with the others, that is, the components must not exhibit the same prediction errors. This lack of correlation is acquired by changing some of the settings in the components, such as the initial conditions, the training data, the network structure or the learning algorithm. Therefore, it is expected that a CN can produce more accurate results than any single expert because knowledge from several experts is combined to reach a decision [13, 28].

The original idea of a committee machine can be traced to Nilsson [23],

who considered a network structure consisting of a layer of elementary perceptrons followed by a voted perceptron scheme in the second layer [6, 13]. This methodology has been applied in various fields of science with accurate results [16, 18, 24].

As stated by Haykin [13], CNs are universal approximators and can be classified into two major categories: static structures (ensemble averaging or boosting) or dynamic structures (mixtures of experts or a hierarchical mixture of experts). In this research, the ensemble averaging category of CNs will be used, in which the outputs of different networks (predictors) are linearly combined to produce an overall output. Despite being one of the simplest methods used in regression problems, Bishop [4] demonstrated that the mean square error (m.s.e.) of an ensemble estimator is guaranteed to be less than or equal to the average m.s.e. of each expert considered in the CN.

A major drawback in the use of a CN, or even of only one ANN, in prediction problems is that it is only possible to obtain point predictions. Unfortunately, point predictions do not provide any information regarding the associated uncertainties or indicate their level of reliability. The measures used to evaluate the ANN's or CN's overall performance, such as the m.s.e., cannot be used to determine regions in which the output of the network can be mistaken, e.g., due to uncertainty arising from random errors or due to a low representation of training data in these regions. An alternative method for solving this problem is to compute a full range of plausible values for the predicted value. In other words, credible prediction intervals can be considered, which assign some uncertainty to the point predictions. Some solutions for calculating ANN prediction intervals have been proposed in the literature [7, 15, 30, 31].

ANNs have been applied in poultry science to predict the energetic va-

lues of feedstuffs. Perai et al. [25] used an MLP to predict the true metabolizable energy (TMEn) values of meat and bone meal samples. A group method of data handling-type neural network accurately predicted the TMEn values of feather meal and poultry offal meal on the basis of their chemical composition [2]. Recently, Mariano et al. [21] proposed an MLP model for predicting values of apparent metabolizable energy corrected for nitrogen balance (AMEn) for feedstuffs commonly used in feed formulation for broilers. These authors selected the neural network that presented the best generalization, and they developed a macro in Excel[®] based on those results, which was called the AMEn Excel Calculator. The calculator enables one to obtain feedstuff predictions for new data. However, among the published neural network studies related to predictions of the metabolizable energy (ME) of poultry feed, none have considered the uncertainty of the predictions, nor has a CN been considered for ME prediction.

The main objective of this study was to associate some uncertainty over the predictions of neural networks models, considering the credibility intervals of predictions of an implemented network committee to predict the AMEn values of the energetic and protein concentrate feeds, generally used by the poultry feed industry. With the objective of quantitatively comparing CNs with different sizes, the percentage of correct predictions (PCP) was used by considering the highest probability density (HPD) intervals for each feedstuff in the test data. In addition, this study sought to determine which of three estimators of the empirical distribution obtained by using the CN ensures the greatest accuracy of prediction: the mean, the median or the mode. In this case, the estimates obtained for the test data were compared with results obtained from metabolic bioassays.

2 Material and methods

2.1 Data source

The database used in the development of each MLP was previously catalogued and described by Nascimento et al. [22]. The database included a total of 568 experimental results corresponding to 370 energetic concentrate and 198 protein concentrate samples evaluated in experiments carried out in Brazil between 1967 and 2007. All feedstuffs were of plant origin. The AMEn, the chemical composition values (crude protein - CP; ether extract without acid hydrolysis - EE; ash; crude fiber - CF), the classification of the feedstuff category (1 - energetic concentrate; 2 - protein concentrate), the specification of the ingredient (1 - maize or soybean meal; 2 - maize byproduct or other soybean by-product; 3 - other feedstuff) and the type of animal used in the bioassay (1 - chick; 2 - rooster) were defined for each feedstuff. The AMEn was determined on the basis of a bioassay. The NDF (neutral detergent fiber) and ADF (acid detergent fiber) values were not considered because this information was not available for the vast majority of experimental results.

Test data were assembled to evaluate the proposed model. These data were obtained from metabolic bioassays conducted in the Poultry Section of the Department of Animal Science at the Federal University of Lavras. Some of these experimental results were described by Alvarenga et al. [3], and the remainder is under review for publication. These data refer to 48 samples of energetic and protein feedstuffs. The AMEn, CP, EE, Ash and CF values were determined for each feedstuff. Furthermore, the feedstuff category, type of ingredient, type of animal and observed AMEn were determined.

2.2 Model development

The AMEn values of concentrate feedstuffs were predicted by a CN. Specifically, several MLPs were implemented using meta-analysis in each development network. Meta-analysis is a relevant method for summarizing and quantifying knowledge acquired through previously published experimental results; that is, new results can be obtained from the synthesis of n distinct but related observations [19, 29]. Thus, each MLP developed from a meta-analysis is considered as an expert network and is a candidate for inclusion in the CN.

The MLP structure was defined as 7-p-q-1, which represents seven inputs, p neurons in the first hidden layer, q neurons in the second hidden layer and one output. Two hidden layers were chosen in the development of the MLP because this structure allows for any function approximation [8]. The input parameters of the implemented MLP were feedstuff categories (CAT), concentrate ingredients (ING), type of animals used in the bioassay (ANIMAL), CP, EE, MM and CF. Different values for p and q were considered, with the objective of identifying the structures that present the best performance. Values from 1 to 300 were evaluated for both p and q .

Two supervised learning algorithms were used to train the network and to adjust the parameters (synaptic weights and bias) of the MLP: the Levenberg-Marquardt (LM) algorithm, described by Bishop [4], and the Resilient backpropagation (RPROP) approach described by Riedmiller and Braun [27]. The purpose of these algorithms is to minimize the m.s.e. of the prediction, which quantifies the difference between the desired outcome (observed AMEn) and the predicted response obtained by the output neuron.

The data were standardized on the interval $[-1, 1]$ before being presented to the network training [13]. The outputs obtained from the trained ANN were then

returned to the original scale. The activation functions were a hyperbolic tangent (sigmoid function) for the hidden layers and a linear function for the output layer.

The dataset used to adjust each MLP was randomly partitioned into two separate subsets: one comprised 85% (483 studies) of the data and was used as the training set in the development of the network; the second, comprising 15% (85 studies) of the data, was used as a validation set. The training set was used to estimate the synaptic weights and bias of the model, and the validation set was used to simultaneously verify the adjustment and the generalization of the network. Two criteria were used to stop the network training: the early-stopping criterion, based on an acceptable m.s.e., and a maximum of 1000 training epochs.

The test dataset, obtained from bioassays, was used to verify the generalization ability of each developed network. Note that this dataset was not used in either training or validation. The performance of the predictions was verified based on the values of the m.s.e., as defined below. Lower m.s.e. values indicate a more accurate MLP model.

Two strategies were used to construct the CN. In one strategy, the CN was composed of the ten different MLP structures that presented the best performance, independent of the learning algorithm used. The MLPs with the best performance were those networks that presented the lowest m.s.e. for the test data and, consequently, the best generalization capacity.

In the other strategy, the MLP structure that presented the best performance was identified, and 1000 MLPs were trained, validated and tested, keeping the optimal structure and using the same learning algorithm. The initial weights for each of the 1000 MLPs were different, and the training and validation datasets were randomly defined for each. Thus, as stated by Lima [17], each of the 1000 networks reached a different generalization, even when considering the same

network structure, because the weights and biases were initialized with random values. Each one of these 1000 MLPs, called expert network, is a candidate to be considered by the committee used to predict the AMEn values of poultry feeds.

After creating the 1000 expert networks, the top N networks, with $N = 10, 11, 12, \dots, 1000$, that presented the best performance were selected to form a CN. The performance criterion was the m.s.e., with the lowest being the best. Thus, the AMEn predictions were based on the results of three or more networks instead of only one. A better committee generalization is obtained with N expert networks that have a higher ability to generalize [17]. Fig. 2 illustrates the structure of a committee using an ensemble of MLP models that share common inputs for training and whose N individual outputs (y_i) are somehow combined to produce an overall output. The expert networks are identified by MLP i , with $i = 1, \dots, N$. The CN was then evaluated using the same test dataset employed for testing individual networks.

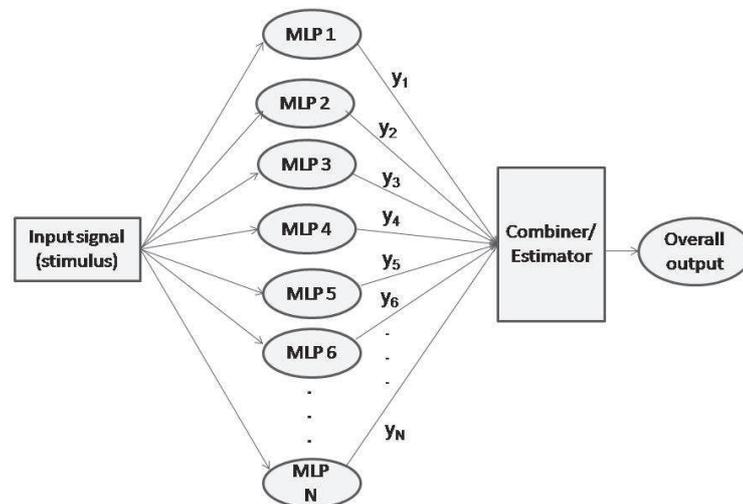


Fig. 2 Structure of a committee with N neural networks

Different numbers of expert networks ($N = 10, 11, \dots, 1000$) were considered to identify the most appropriate number for predicting AMEn values. The N outputs were combined with the objective of obtaining an overall predicted value of the AMEn. First, the empirical distribution of the predicted values of the AMEn was obtained with the N predicted values. Then, three estimators were considered to represent the overall predicted value: the mean, the median and the mode of the empirical distribution. The objective was to define which of these estimators best represented the overall predicted value of the AMEn.

The PCP was used to evaluate and compare the prediction capacity of each CN. Credibility intervals were used to quantify the PCP. The credibility interval for each feedstuff in the test dataset was obtained as follows:

- an empirical distribution was constructed for the N predicted values of the AMEn for a specific feedstuff;
- from the empirical distribution, a range was built that included the points of maximum density and minimum length, which represents the HPD (highest probability density) interval. This range was considered to have a credibility level of 95%.

These HPD intervals take into account some uncertainty over the estimates of interest. Therefore, these intervals were used to verify whether the committee ensured accurate predictions compared to the AMEn values observed in the bioassays. If the observed value was included in the HPD, there is no evidence that the observed value differs of the predicted value.

Thus, the PCP was obtained from the frequency of the observed values of

the feedstuffs within the respective intervals. This percentage can be expressed as

$$PCP = \frac{\sum_{i=1}^K I(LL_i \leq Y_{OBS_i} \leq LS_i)}{K} \cdot 100, \quad (0.1)$$

where:

$I(\cdot)$ is an indicator function, with $I(\cdot) = \begin{cases} 1, & \text{if } Y_{OBS_i} \in \text{HPD} \\ 0, & \text{otherwise} \end{cases}$;

LL_i and LS_i are the lower and upper limits, respectively, of the HPD interval for the i th feedstuff;

Y_{OBS_i} is the observed value of the AMEn for the i th feedstuff;

K is the number of feedstuffs in the test dataset.

The performance and accuracy of the committees were evaluated with the three estimators by using the following measures, as defined in Perai et al. [25]:

- R^2 (coefficient of determination)

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$$

- m.s.e. (mean square error)

$$m.s.e. = \frac{\sum_{i=1}^n |y_i - \hat{y}_i|^2}{n}$$

- MAD (mean absolute deviation)

$$MAD = \frac{\sum_{i=1}^n |y_t - \hat{y}_t|}{n}$$

- MAPE (mean absolute percentage error)

$$MAPE = \frac{\sum_{i=1}^n \left| \frac{y_t - \hat{y}_t}{y_t} \right|}{n} \cdot 100, (y_t \neq 0)$$

- Bias

$$Bias = \frac{\sum_{i=1}^n y_t - \hat{y}_t}{n}$$

where y_t is the observed value, \hat{y}_t is the predicted value, \bar{y}_t is the mean of the observed values and n is the number of observations.

The software MATLAB [20] was used to implement the networks because it offers a very efficient ANN toolbox. All other statistical analyses were performed with the R software [26].

3 Results and discussion

Of the different structures initially tested, the 7-5-3-1 MLP models were more accurate in predicting the AMEn values of poultry feeds because they ensured the lowest m.s.e. values, considering the test data. The best performance was obtained with the LM learning algorithm. This algorithm is a modification of the Gauss-Newton algorithm and has a higher convergence speed [11]. This feature finding does not mean that the LM is always better than other algorithms, but it proved to

be the most appropriate for the situation proposed in this paper.

The training process of all networks was terminated by the early-stopping criterion, i.e., before 1000 epochs. This high limit number (1000 epochs) was chosen to ensure that the training was stopped before the limit was reached so that the m.s.e. in the validation data did not grow and then decrease. The performance of the committees with different numbers of networks (N) was assessed with the PCP metric; the results are presented in Table 1.

Table 1 Percentage of correct predictions (PCP) for the test data using various committee sizes (N) with the same structure

	Number of expert networks (N)							
	10	50	100	300	500	700	900	1000
PCP (%)	60.4	66.7	68.8	70.8	72.9	75.0	79.2	95.8

The highest PCP value (PCP=95.8%) was obtained with N = 1000 expert networks in the committee (Table 1). This committee gave the most accurate predictions; it correctly predicted more than 95% of the 48 feedstuffs in the test dataset. This result is quite satisfactory as it implies that the observed AMEn value of 46 feedstuffs was statistically equal to the committee-predicted value. It is important to highlight that the smaller the N, the lower the PCP. Therefore, if a PCP of 70% is considered satisfactory, a committee with at least 300 networks is needed. This result contradicts those of other studies in the literature, in which the use of a few networks (on average, N = 15) usually provides excellent committee performance [18, 24].

Some authors have used CNs composed of expert networks with different structures and even different learning algorithms [16, 17, 24]. The advantage of this kind of CN is that it avoids correlated components [17]. In this work, a CN was

Table 2 Structures and learning algorithms of the ten expert networks selected to construct a CN

Expert networks	MLP structure	Learning algorithm
1	7-7-7-1	RPROP
2	7-8-5-1	RPROP
3	7-5-3-1	RPROP
4	7-7-7-1	LM
5	7-8-5-1	LM
6	7-8-6-1	LM
7	7-8-7-1	LM
8	7-5-4-1	LM
9	7-5-5-1	LM
10	7-5-3-1	LM

LM - Levenberg-Marquardt; RPROP - resilient backpropagation.

formed with the ten MLP structures that presented the best performance (Table 2). However, the PCP of this CN was only 52.1%.

Better prediction results were expected for this last committee because it was intended to eliminate the possible correlations that may exist among expert results [17]. Its PCP was less than the PCP of the CN with ten MLPs with the same structure (Table 1). However, in these cases, where the number of expert networks (N) in a CN is small, the approximation of the empirical distribution is harmed, and the PCP values present low reliability. Thus, a committee of 1000 expert MLPs (Table 1) is recommended for predicting the AMEn values of poultry feeds.

The empirical AMEn distributions for each of the 48 feedstuffs (test dataset) were obtained from N predicted values of the CN. This finding illustrates a great advantage in using a CN to estimate any variable because one can engender a distribution of probabilities associated with predicted values. Consequently, inferences can be made, and the degree of uncertainty of the prediction can be measured.

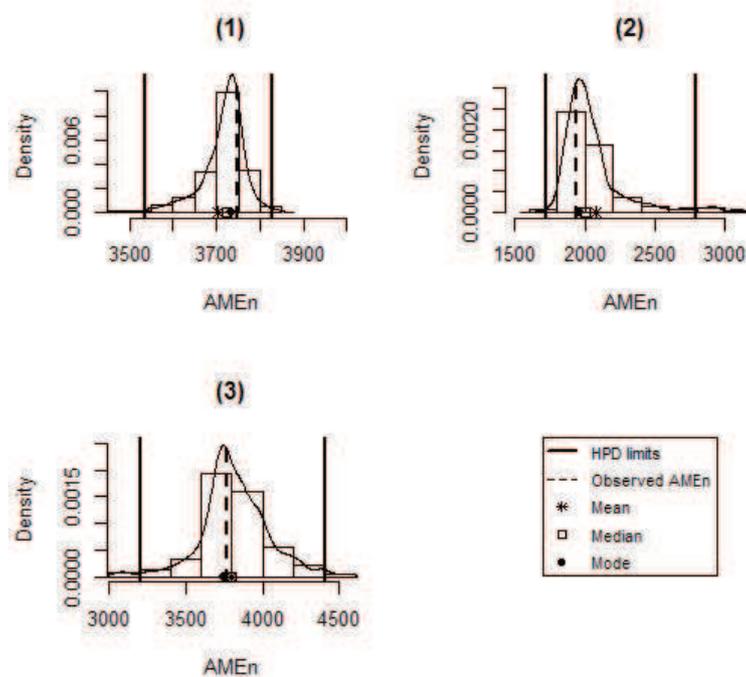


Fig. 3 Empirical distributions, observed and predicted values of the AMEn and respective HPD interval limits of three feedstuffs in the test dataset

As an example, the empirical distributions for three feedstuffs in the test data using $N = 1000$ predicted values are shown in Fig. 3. The observed AMEn value, the predictions obtained by the estimators (the mean, the median and the mode of the empirical distributions) and the limits of the HPD intervals are highlighted.

The 48 empirical distributions for the predicted AMEn values were not similar. There were symmetric and asymmetric distributions, with different forms and degrees of skewness and kurtosis, as can be observed in Fig. 3. Most distributions were asymmetric.

Because one of the goals was to estimate the AMEn from these distributions, it is necessary to concisely express the true value of the feedstuff AMEn

values using summary measures such as the mean, median and mode. Several works, such as Lima [17], Lipnickas [18] and Siwek et al. [31], used the mean to summarize the information generated by each expert network in a CN. However, if the distribution is asymmetric, there are extreme values greater or smaller than the mean in only one direction, and the mean is greatly influenced by these extremes. Therefore, for asymmetric distributions, the mean is a poorer estimator than the median and mode.

The estimated and observed values of the AMEn, as well as the HPD interval limits obtained from the empirical distribution in Fig. 3, are given in Table 3. It can be seen that the estimates of the mean are less accurate than those of the median and the mode for the three examples. This behavior was observed for the majority of the feedstuffs in the test dataset. The exceptions are cases in which the empirical distributions are more symmetric, where, as expected, the estimates of the mean, median and mode were similar.

Table 3 Observed AMEn and estimates of the mean, median, mode and HPD interval limits from the empirical distributions of three feedstuffs in the test dataset

Observed AMEn	Mean	Median	Mode	HPD 95%
3747	3701.9	3724.5	3735.8	[3537; 3828]
1941	2079.4	2000.0	1957.5	[1722; 2791]
3772	3799.4	3795.0	3744.5	[3204; 4407]

In this study, HPD intervals with 95% credibility were used to determine whether the predictions were accurate. As shown in Table 3, the three HPD intervals included the observed or true value of the AMEn. Therefore, the predictions were accurate. This trend was also observed for the remaining 45 feedstuffs in the test dataset. In practice, when the true value of the AMEn is unknown, an HPD

interval with 95% credibility indicates a 95% probability that the interval includes the true value.

A summary of the statistical results associated with the CN with 1000 networks, considering the different estimators, is shown in Table 4. These statistics measure the prediction errors based on the difference between the observed and predicted values for the feedstuffs in the test dataset. As the R^2 value approaches one, the Bias value approaches zero, and lower m.s.e., MAPE and MADE values indicate a more accurate model. These statistics were calculated based on 48 feedstuffs in the test dataset and are needed to evaluate the performance of each estimator, considering estimates obtained from the empirical distributions.

Table 4 Statistics for evaluating the accuracy and goodness-of-fit of the mean, median and mode estimators, considering estimates obtained from the empirical distribution of the CN with 1000 networks

Statistics	Estimators		
	Mean	Median	Mode
R^2	0.88	0.89	0.89
m.s.e.	50897.15	45889.71	45285.43
MAD	193.40	177.58	175.66
MAPE (%)	6.81	6.23	5.97
Bias	-84.34	-102.46	-86.80

R^2 - coefficient of determination; MAD - mean absolute deviation; MAPE - mean absolute percentage error; m.s.e. - mean square error.

R^2 values greater than 88% were obtained for all three estimators. This finding indicates a high performance of the prediction models for all estimators, showing that all estimators lead to estimates close to the true values (Table 4). However, the mode of the empirical distribution obtained lower m.s.e., MAD and MAPE values in comparison to the mean and the median, although the mean has been used by some researchers [17, 18, 31].

Based on these results, the mode appears to be the most appropriate estimator. However, the mode is calculated using different approximation methods. Because approximated calculations are involved, care must be taken when using the mode. Therefore, the median may be a better option for estimating the AMEn from an empirical distribution.

4 Conclusions

The use of a network committee to predict the AMEn values of poultry feeds is a novel and unpublished approach that leads to precise and accurate estimations. The use of a network committee is recommended because it allows one to associate some uncertainty with the prediction using an HPD interval.

The mode of the empirical distribution obtained from the committee, rather than the mean, is recommended for predicting the AMEn of energetic and protein feedstuffs for broilers. However, care is required when choosing an approach for calculating the mode because the calculations include approximations. If a researcher is uncertain about which method to use, the median is recommended. The important point is that both the mode and the median are more suitable than the mean.

The committee consisting of $N = 1000$ networks had the highest percentage of correct predictions (PCP=95.8%). However, even the committee of $N = 300$ networks had success rate above 70%. Therefore, the use of at least 300 networks is recommended in a committee of networks, which is far greater than the number of networks generally used in the literature.

Acknowledgements

The authors gratefully acknowledge financial support provided by the Coordination Improvement of Higher Education Students (CAPES).

References

1. Ahmadi H, Mottaghitalab M, Nariman-Zadeh N (2007) Group method of data handling-type neural network prediction of broiler performance based on dietary metabolizable energy, methionine, and lysine. *Journal of Applied Poultry Research* 16: 494-501. doi: 10.3382/japr.2006-00074.
2. Ahmadi H, Golian A, Mottaghitalab M, Nariman-Zadeh N (2008) Prediction model for true metabolizable energy of feather meal and poultry offal meal using group method of data handling-type neural network. *Poultry Science* 87: 1909-1912. doi: 10.3382/ps.2007-00507.
3. Alvarenga RR, Rodrigues PB, Zangeronimo MG, Freitas RTF, Lima RR, Bertechini AG, Fassani EJ (2011) Energetic values of feedstuffs for broilers determined with in vivo assays and prediction equations. *Animal Feed Science and Technology* 168: 257-266. doi: 10.1016/j.anifeedsci.2011.04.092.
4. Bishop CM (1995) *Neural networks for pattern recognition*. Oxford University Press, Oxford, UK.
5. Bucene LC, Rodrigues LHA (2004) Utilização de redes neurais artificiais para avaliação de produtividade do solo, visando classificação de terras para irrigação. *Revista Brasileira de Engenharia Agrícola e Ambiental* 8: 326-329. doi: 10.1590/S1415-43662004000200025.

6. Can M (2013) Committee Machine Networks to Diagnose Cardiovascular Diseases. *Southeast Europe Journal of Soft Computing* 2: 76-83.
7. Chryssolouris G, Lee M, Ramsey A (1996) Confidence interval prediction for neural network models. *IEEE Trans. Neural Networks* 7: 229-232. doi: 10.1109/72.478409.
8. Cybenko G (1988) Continuous valued neural network with two hidden layers are sufficient. Technical Report, Department of Computer Science, Tufts University, Medford, MA, USA.
9. da Silva IN, Spatti DH, Flauzino (2010) *Rede Neurais Artificiais: Para Engenharia e Ciências Aplicadas*. Ed. Artliber, São Paulo.
10. Goldschmidt RR (2010) *Uma Introdução à Inteligência Computacional: fundamentos, ferramentas e aplicações*. 1ed. Rio de Janeiro: IST-Rio.
11. Hagan MT, Menhaj MB (1994) Training feedforward networks with the Marquardt algorithm. *IEEE Transactions on Neural Networks* 5: 989-993. doi: 10.1109/72.329697.
12. Haider A, Hanif MN (2009) Inflation forecasting in Pakistan using artificial neural networks. *Pakistan economic and social review* 47: 123-138.
13. Haykin S (2007) *Neural networks - a comprehensive foundation*. 3rd edition. Prentice-Hall Inc., Upper Saddle River, NJ, USA.
14. Hunter D, Wilamowski B (2011) Parallel multi-layer neural network architecture with improved efficiency. *International Conference on Human System Interaction (HSI 2011)*, Yokohama, Japan.

15. Hwang JTG, Ding AA (1997) Prediction intervals for artificial neural networks. *J. American Statistical Association* 92: 748-757. doi: 10.1080/01621459.1997.10474027.
16. Kenari SAJ, Mashohor S (2013) Robust committee machine for water saturation prediction. *Journal of Petroleum Science and Engineering* 104: 1-10. doi: 10.1016/j.petrol.2013.03.009.
17. Lima CAM (2004) Comitê de Máquinas: Uma Abordagem Unificada Empregando Máquinas de Vetores-Suporte. Doctoral Thesis, Universidade Estadual de Campinas.
18. Lipnickas A (2008) Adaptive Committees of Neural Classifiers. *Information Technology And Control* 37: 205 - 210.
19. Lovatto PA, Lehnen CR, Andretta I, Carvalho AD Hauschild L (2007) Meta-analysis in scientific research: a methodological approach. *Brazilian Journal of Animal Science* 36 (suppl.): 285-294. doi: 10.1590/S1516-35982007001000026.
20. MATLAB 7.12 R2011a (2011) The Math Works, Inc. Available in: <<http://www.mathworks.com/products/neuralnet/>>. Access on: 25 ago. 2013.
21. Mariano FCMQ, Paixão CA, Lima RR, Alvarenga RR, Rodrigues PB, Nascimento GAJ (2013) Prediction of the energy values of feedstuffs for broilers using meta-analysis and neural networks. *Animal* 7: 1440-1445. doi:10.1017/S1751731113000712.

22. Nascimento GAJ, Rodrigues PB, Freitas RTF, Bertechini AG, Lima RR, Pucci LEA (2009) Prediction equations to estimate the energy values of plant origin concentrate feeds for poultry utilizing the meta-analysis. *Brazilian Journal of Animal Science* 38: 1265-1271. doi: 10.1590/S1516-35982009000700015.
23. Nilsson NJ (1965) *Learning Machines: Foundations of Trainable Pattern-Classifying Systems*. New York: Macgraw-Hill.
24. Oliveira RC, Acevedo NIA, Silva Neto AJ, Biondi Neto L (2010) Aplicação de um comitê de redes neurais artificiais para a solução de problemas inversos em Transferência Radiativa. *TEMA - Tend. Mat. Apl. Comput.* 11: 171-182. doi: 10.5540/tema.2010.011.02.0171.
25. Perai AH, Moghaddam HN, Asadpour S, Bahrampour J, Mansoori GH (2010) A comparison of artificial neural networks with other statistical approaches for the prediction of true metabolizable energy of meat and bone meal. *Poultry Science* 89: 1562-1568. doi: 10.3382/ps.2010-00639.
26. R DEVELOPMENT CORE TEAM (2013) *R: a language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing. Available in: <<http://www.r-project.org>>. Access on: 10 oct. 2013.
27. Riedmiller M, Braun H (1993) A direct adaptive method for faster backpropagation learning: The RPROP algorithm. *Proceedings of the IEEE International Conference on Neural Networks (ICNN)*: 586 - 591, San Francisco.

28. Rustempasic I, Can M (2013) Diagnosis of Parkinson's Disease using Principal Component Analysis and Boosting Committee Machines. SOUTHEAST Europe Journal of Soft Computing 2: 102-109.
29. Sauvant D, Schmidely P, Daudin JJ, St-Pierre NR (2008) Meta-analyses of experimental data in animal nutrition. *Animal* 2: 1203-1214. doi: 10.1017/S1751731108002280.
30. Shao R, Martin EB, Zhang J, Morris AJ (1997) Confidence bounds for neural network representations. *Computers & Chemical Engineering* 21(suppl.): S1173-S1178. doi: 10.1016/S0098-1354(97)87661-2.
31. Siwek K, Osowski S, Szupiluk R (2009) Ensemble neural network approach for accurate load forecasting in a power system. *International Journal of Applied Mathematics and Computer Science* 19: 303-315. doi: 10.2478/v10006-009-0026-2.

ARTIGO 3: Comitê de redes neurais artificiais e regressão múltipla ponderada para a predição de valores energéticos de alimentos para aves de corte.

Redigido conforme a norma para publicação periódica científica NBR 6022 da ABNT.

F.C.M.Q. Mariano^{1,3}, R.R. Lima¹, R.R. Alvarenga² e P.B. Rodrigues²

¹Departamento de Ciências Exatas, Universidade Federal de Lavras, Cx. Postal 3037, 37200-000 Lavras, Minas Gerais, Brasil. ²Departamento de Zootecnia, Universidade Federal de Lavras, Cx. Postal 3037, 37200-000 Lavras, Minas Gerais, Brasil. ³Email para correspondência: flaviaqz@gmail.com

RESUMO

Este trabalho foi realizado com o objetivo de comparar duas metodologias anteriormente propostas para a predição de energia metabolizável aparente corrigida pelo balanço de nitrogênio (EMAn) de alimentos para aves de corte. Foram comparados um modelo de comitê de redes neurais e uma equação de predição ajustada por modelo de regressão linear múltipla ponderada (RLMP), considerando o princípio da meta-análise e utilizando a técnica de componentes principais na formação de grupos homogêneos de resultados experimentais. Para essa comparação, foram consideradas porcentagens de acerto de predição, com base em intervalos de predição e intervalos de credibilidade de máxima densidade de probabilidade (HPD) e, em seguida, foi utilizado um teste para comparação de proporções. Além disso, a acurácia de cada modelo selecionado foi avaliada com base nos valores de m.s.e. (erro quadrado médio), R^2 , MAD (desvio médio absoluto), MAPE (erro médio percentual absoluto) e Bias. Dados provenientes de ensaios metabólicos foram utilizados como conjunto de teste na comparação dos modelos. Os resultados revelaram que o comitê de redes produziu predições mais acura-

das de EMAn, sendo indicado como um melhor método para a rápida predição do valor energético de alimentos concentrados energéticos e proteicos comumente utilizados na formulação de rações avícolas.

Palavras-chave: energia metabolizável, intervalo HPD, meta-análise, porcentagem de acerto, rações avícolas.

ABSTRACT

The purpose of this study was to compare previously developed methods to estimate the nitrogen-corrected apparent metabolizable energy (AMEn) of poultry feedstuffs. A network committee model and a prediction equation adjusted by using a weighted multiple linear regression and meta-analysis principle were compared. The comparison of these models was performed by considering the percentages of correct prediction, based on the classic prediction intervals and the highest probability density (HPD) intervals, and by using comparison test for proportions. Furthermore, the accuracy of models was evaluated using the mean squared error, R^2 , the mean absolute deviation, the mean absolute percentage error and the Bias. Data from metabolic trials were used to compare the selected models. The model that provided the highest accuracy of prediction was the network committee model. Thus, it was recommended this method as a rapid and efficient way to predict the AMEn values of concentrate feedstuffs for broilers.

Keywords: HPD interval, meta-analysis, metabolizable energy, percentage of success, poultry feedstuffs.

1 INTRODUÇÃO

Em razão da diversidade de alimentos e seus subprodutos, utilizados na formulação de rações para aves, surge a necessidade de um conhecimento preciso

da composição química e dos valores de energia metabolizável (EM) dos alimentos, o que possibilita um fornecimento adequado de energia para os animais. A determinação da composição química é bem mais fácil e menos onerosa do que determinar a energia metabolizável dos alimentos. Uma das formas mais diretas de determinar a energia metabolizável é utilizar equações de predição.

Métodos estatísticos clássicos têm sido aplicados para a obtenção de equações de predição da energia metabolizável aparente corrigida pelo balanço de nitrogênio (EMAn) de alimentos para aves. Nestas equações, os valores energéticos dos alimentos são estabelecidos em função da composição química dos mesmos. Nos métodos empregados na obtenção dessas equações de predição são consideradas diferentes formas de se desenvolver uma regressão linear múltipla: utilizando-se a análise de regressão múltipla clássica, sem qualquer tipo de restrição ou ponderação (RODRIGUES et al., 2002); ou utilizando-se o princípio da meta-análise, em que grupos homogêneos são estabelecidos como fatores de ponderação (MARIANO et al., 2012; NASCIMENTO et al., 2009).

Modelos de redes neurais artificiais (ANN) têm sido utilizados, recentemente, na predição de valores energéticos de alimentos para aves (MARIANO et al., 2013b; PERAI et al., 2010). Devido à sua grande capacidade de aprender por meio de exemplos, generalizar a informação aprendida e, posteriormente, dar respostas coerentes para dados não conhecidos, é possível afirmar que a aplicação de uma ANN vai além do que simplesmente mapear relações de entrada e saída. Por este fato, os métodos estatísticos e os métodos de ANN têm competido entre si por espaço na construção de modelos em diversas áreas (ANYAECHE; IGHRAVWE, 2013; SHI et al., 2012). Modelos de ANN, também, têm demonstrado desempenho superior aos métodos de regressão na área de produção e nutrição animal (AHMADI; GOLIAN, 2010; PERAI et al., 2010; WANG; CHEN; ROAN, 2012).

No entanto, para a obtenção de resultados ainda mais acurados em modelos de ANN, recomenda-se a utilização de um comitê de redes. Um comitê de redes representa a agregação de mais de um componente ou rede para se chegar a uma solução global que é supostamente superior àquela obtida por qualquer um dos componentes isolados (HAYKIN, 2007). O objetivo de um comitê de redes é aumentar a capacidade de generalização de estimadores, evitando o sobreajuste dos dados e a convergência da rede para um mínimo local.

O objetivo geral deste estudo foi o de avaliar o desempenho de predição de um modelo de comitê de redes neurais comparado com um modelo de regressão linear múltipla ponderada (RLMP). Estes modelos foram obtidos para a rápida predição de valores de EMAn de diversos alimentos concentrados comumente utilizados em rações para aves de corte.

2 MATERIAL E MÉTODOS

2.1 Conjunto de dados

O conjunto de dados utilizado, para o ajuste do modelo RLMP, foi descrito, anteriormente, em Nascimento et al (2009). Estes dados referentes a 568 trabalhos experimentais, correspondem a 370 amostras de alimentos concentrados energéticos e 198 de concentrados proteicos, conduzidos no Brasil desde 1967 até 2007.

O conjunto de dados utilizado na comparação dos modelos é chamado de conjunto de teste. Estes dados referem-se a 48 amostras de alimentos concentrados energéticos e proteicos, os quais foram determinados em ensaios metabólicos conduzidos no Setor de Avicultura do Departamento de Zootecnia da Universidade Federal de Lavras. Dentre estes dados, encontram-se os resultados experimentais descritos em Alvarenga et al. (2011) e o restante está sob revisão para publica-

ção. Os alimentos proteicos utilizados nestes ensaios metabólicos foram amostras de seis marcas comerciais de farelos de soja, uma de farelo de soja semi-integral, uma de soja integral extrusada, uma de proteína texturizada de soja, duas de soja integral micronizada, duas de farelo de glúten de milho 60% e uma de soja integral tostada. Já os alimentos energéticos compreenderam amostras de quatro híbridos de milho, três de sorgo, uma de milho pré-gelatinizado, uma de quirera de milho, uma de farelo de trigo, uma de farelo integral de arroz, uma de quirera de arroz e uma de gérmen de milho desengordurado.

Os valores observados de EMAn, de composição química (PB - proteína bruta; EE - extrato etéreo; MM - matéria mineral; FB - fibra bruta), a classificação da categoria do alimento (1 - concentrado energético; 2 - concentrado proteico), o tipo de ingrediente (1 - milho ou farelo de soja; 2 - subproduto do milho ou outro subproduto da soja; 3 - outro alimento) e o tipo de animal usado no bioensaio (1 - pinto; 2 - galo) foram definidos para cada alimento. Todos os alimentos considerados são de origem vegetal e são comumente utilizados em rações avícolas.

2.2 Modelos selecionados

2.2.1 Comitê de redes neurais

Neste trabalho, foi considerado um comitê de redes neurais multilayer perceptron (MLP), o qual fora desenvolvido para prever a EMAn de alimentos concentrados energéticos e proteicos para aves e é apresentado por Mariano et al. (2013a). Todas as redes consideradas no comitê possuíam a mesma arquitetura MLP 7-5-3-1, que representa: sete entradas, 5 neurônios na primeira camada oculta, 3 neurônios na segunda camada oculta e uma saída. Foi utilizado o mesmo algoritmo de treinamento, Levenberg-Marquardt, em todos os ajustes, porém com os pesos iniciais e a divisão entre os dados de treinamento e de validação alea-

toriamente definidos pelo software. Após serem testados comitês com diferentes números de redes, o comitê com N=1000 redes garantiu as predições mais acuradas, sendo o estimador da moda da distribuição empírica utilizado como o combinador das redes especialistas. A escolha deste comitê foi baseada no valor da porcentagem de acerto da predição de 95,83%. Outras informações sobre o desenvolvimento deste comitê podem ser obtidos em Mariano et al. (2013a).

2.2.2 Modelo de regressão linear múltipla ponderada (RLMP)

Mariano et al. (2012) propuseram e avaliaram algumas equações de predição da EMAN de alimentos para aves, com base em sua composição química. Estas equações de predição foram ajustadas por modelos RLMP, considerando o princípio da meta-análise, sendo a técnica de componentes principais utilizada na formação de grupos homogêneos de resultados experimentais. Neste estudo, seguindo a metodologia de Mariano et al. (2012) e utilizando o conjunto de 568 resultados experimentais, foi ajustada uma equação de regressão linear múltipla ponderada, considerando os efeitos principais e de interações entre as variáveis explicativas (EqPred).

2.3 Critério de verificação da acurácia dos modelos selecionados

2.3.1 Porcentagem de acerto da predição (PCP)

Para efeito de comparação dos modelos considerados neste trabalho, utilizou-se a porcentagem de acerto da predição (percentage of correct predictions - PCP) no conjunto de dados de teste. Essa porcentagem foi obtida, considerando a frequência em que, no conjunto de teste, os valores observados (valores reais) dos alimentos se encontravam dentro dos seus respectivos intervalos de predição ou de máxima densidade de probabilidade (HPD - *highest probability density*).

Para o modelo de RLMP, calculou-se o intervalo de confiança de predição (FERREIRA, 2005) para cada alimento do conjunto de teste. Já no caso do modelo de comitê de redes, considerou-se o intervalo de credibilidade de predições HPD, pois a distribuição empírica das predições advindas das redes neurais que compunham o comitê era assimétrica (BONAT et al., 2012). O nível de confiança e/ou credibilidade considerado foi de 95%.

Os passos para a obtenção do intervalo de credibilidade HPD para cada alimento do conjunto de teste foram: (i) considerar as predições correspondentes às N=1000 redes do comitê; (ii) obter a distribuição empírica das 1000 predições do alimento; (iii) obter, a partir da distribuição empírica, o intervalo de menor comprimento, que inclui os pontos de máxima densidade, ou seja, o HPD.

Dessa forma, a PCP pode ser expressa pela seguinte razão:

$$PCP = \frac{\sum_{i=1}^K I(LI_i \leq Y_{OBS_i} \leq LS_i)}{K} \cdot 100, \quad (1)$$

em que:

$I(.)$ é uma função indicadora, com $I(.) = \begin{cases} 1, & \text{se } Y_{OBS_i} \in \text{intervalo} \\ 0, & \text{caso contrário} \end{cases}$;

LI_i e LS_i são, respectivamente, os limites inferior e superior do intervalo de predição ou de credibilidade HPD, calculado para o i-ésimo alimento;

Y_{OBS_i} é o valor de EMAn observado (valor real) para o i-ésimo alimento;

K é o número de alimentos contidos no conjunto de teste.

Em (1), diz-se que não há evidências de que o valor predito seja diferente do observado quando o valor observado de um alimento se encontrar dentro do seu respectivo intervalo, resultando no sucesso deste evento. Assim, o cálculo da PCP nada mais é que uma média de uma amostra aleatória simples de uma população

com distribuição de Bernoulli.

2.3.2 Critério para comparação de proporções

Para realizar o teste para comparação das proporções obtidas na PCP, foi utilizado o teste assintótico χ^2 de Pearson, como definido em Biase e Ferreira (2009). A hipótese de nulidade deste teste é a igualdade das proporções binomiais, $H_0 : PCP_1 = PCP_2$, em que PCP_i é o estimador da proporção de acerto de cada um dos modelos selecionados. Ou seja, PCP_1 e PCP_2 referem-se, respectivamente, às porcentagens de acertos com o uso da equação de predição EqPred e do modelo de comitê de redes. Com este teste, é possível verificar se as proporções não diferem entre si. A estatística deste teste possui distribuição qui-quadrado assintótica com $\nu = k - 1$ graus de liberdade, sendo k o número de populações (neste estudo, o número de modelos selecionados). O nível de significância adotado para o teste de comparação das proporções foi de 5%.

2.3.3 Critérios de adequabilidade de modelos

A qualidade do ajuste e a acurácia dos modelos de predição da EMAn, tanto da equação de predição quanto do comitê de redes, foram avaliados com os dados de teste. As estatísticas utilizadas neste processo, baseadas nos erros entre os valores de EMAn observados e preditos, foram as seguintes: R^2 (coeficiente de determinação), m.s.e. (erro quadrático médio), MAD (desvio médio absoluto), MAPE (erro médio percentual absoluto) e Bias, conforme definido em Mariano et al. (2013a) e Perai et al. (2010). Assim, quanto mais próximo de um for o valor de R^2 , mais próximo de zero for o valor do Bias, e menores os valores de m.s.e., de MAPE e de MAD, mais acurado é o modelo.

Todas as análises estatísticas foram realizadas no software livre R (R DE-

VELOPMENT CORE TEAM, 2013).

3 RESULTADOS E DISCUSSÃO

A EqPred foi ajustada em busca de uma melhor capacidade preditiva dentre os métodos estatísticos usuais, considerando os efeitos principais e de interações entre as variáveis explicativas. Para este ajuste, foi criada a variável TRAT por meio da combinação entre os níveis das variáveis categóricas do conjunto de dados (categoria do alimento, tipo de ingrediente e o tipo de animal usado no bi-ensaio). Dessa forma, o ajuste do modelo RLMP se baseia em 12 equações de regressão, em que cada TRAT possui uma equação específica (Tabela 1).

Os modelos selecionados neste estudo, EqPred e Comitê de redes neurais, foram desenvolvidos para a predição da EMAn de alimentos concentrados tanto energéticos quanto proteicos. No entanto, o uso destes modelos é, ainda, mais apropriado para estimar a EMAn de alimentos tais como o milho, sorgo, soja, farelo de soja, trigo e farelo de trigo (MARIANO et al., 2012, 2013b). Isto se deve ao fato de que estes alimentos tiveram uma maior representatividade no conjunto de dados utilizados no desenvolvimento dos modelos.

Para o modelo de comitê de redes, foram considerados intervalos HPD, pois a distribuição empírica das predições advindas das 1000 redes neurais que compunham o comitê era assimétrica. Dessa forma, o melhor intervalo a ser calculado foi o intervalo de credibilidade de predição que incluísse os pontos de máxima densidade e menor comprimento, ou seja, o HPD (BONAT et al., 2012). Para os modelos de equação de predição, foram considerados intervalos de predição (IP), os quais, também, podem ser definidos como intervalos HPD, pois correspondem ao menor intervalo de máxima cobertura devido à distribuição normal simétrica assumida pelo uso de regressão linear múltipla.

Tabela 1 Equações de predição para predizer os valores da EMAn de alimentos para frangos de corte, específicas para cada combinação da variável TRAT

TRAT _{ijk} *	Equação específica
TRAT ₁₁₁	3196,898+111,127EE+69,109MM-28,856FB+41,740PB+ -13,987EE.PB+2,699MM.PB+8,107FB.PB+0,362EE.MM.PB
TRAT ₁₁₂	2605,700+111,127EE+69,109MM-28,856FB+148,109PB+ -8,423EE.PB-62,111MM.PB+8,297FB.PB+2,381EE.MM.PB
TRAT ₁₂₁	2605,067+111,127EE+69,109MM-28,856FB+213,263PB+ -22,078EE.PB-21,278MM.PB-10,972FB.PB+2,888EE.MM.PB
TRAT ₁₂₂	8651,232+111,127EE+69,109MM-28,856FB-755,557PB+ +8,571EE.PB+13,079MM.PB+14,284FB.PB
TRAT ₁₃₁	3272,000+111,127EE+69,109MM-28,856FB+38,562PB+ -7,965EE.PB-5,603MM.PB-6,819FB.PB+0,200EE.MM.PB
TRAT ₁₃₂	3670,277+111,127EE+69,109MM-28,856FB-72,692PB+ +7,042EE.PB+0,046MM.PB-6,248FB.PB-1,088EE.MM.PB
TRAT ₂₁₁	1530,008+111,127EE+69,109MM-28,856FB-16,175PB+ +14,131EE.PB+3,377MM.PB+1,456FB.PB-2,595EE.MM.PB
TRAT ₂₁₂	965,120+111,127EE+69,109MM-28,856FB+79,883PB+ -17,111EE.PB-9,414MM.PB+1,403FB.PB+2,183EE.MM.PB
TRAT ₂₂₁	718,556+111,127EE+69,109MM-28,856FB+17,382PB+ -0,810EE.PB+1,609MM.PB+0,802FB.PB-0,0193EE.MM.PB
TRAT ₂₂₂	-1057,607+111,127EE+69,109MM-28,856FB+33,665PB+ +5,233EE.PB+5,440MM.PB+2,567FB.PB-1,234EE.MM.PB
TRAT ₂₃₁	991,734+111,127EE+69,109MM-28,856FB+42,695PB+ +0,467EE.PB-3,750MM.PB+0,312FB.PB-0,295EE.MM.PB
TRAT ₂₃₂	4586,733+111,127EE+69,109MM-28,856FB+31,285PB+ -4,715EE.PB-24,654MM.PB+3,502FB.PB

* TRAT_{ijk} representa a equação específica quando o alimento for respectivo à i-ésima categoria do alimento, ao j-ésimo tipo de ingrediente e ao k-ésimo tipo de animal, em que:
i = 1 - concentrado energético, 2 - concentrado proteico;
j = 1 - milho ou farelo de soja, 2 - subproduto do milho ou outro subproduto da soja, 3 - outro alimento;
k = 1 - pintos de corte, 2 - galo.

Os valores de EMAn observados em ensaios metabólicos e seus respectivos intervalos referentes aos modelos considerados neste estudo, estão apresentados na Tabela 2. Esses intervalos foram calculados para os 48 alimentos do conjunto de teste.

Na Tabela 2, verifica-se que a amplitude média dos intervalos de credibilidade HPD foi superior comparada à amplitude média dos intervalos de predição. Porém, esta diferença diminuiu à medida que foram consideradas menos redes no comitê, pois os intervalos dos comitês com N=100, 500 e 800 redes apresentaram, respectivamente, uma amplitude média de 423,54 (PCP=68,75%), 528,83 (PCP=72,92%) e 643,63 (PCP=72,92%) (MARIANO et al., 2013a). Portanto, verificou-se uma tendência de que, quanto maior foi a complexidade do modelo, maior foi a amplitude obtida nos intervalos e, conseqüentemente, maior foi o valor obtido para a PCP (Tabelas 2 e 3).

A estimativa dos intervalos de credibilidade HPD permitiu a associação de alguma incerteza aos valores preditos, aumentando a confiabilidade no comitê de redes neurais (Tabela 2). Além disso, os intervalos construídos foram utilizados como limites de controle para verificar a PCP dos modelos selecionados, cujos valores são apresentados na Tabela 3.

A igualdade das proporções binomiais foi verificada pelo teste assintótico de χ^2 de Pearson, o qual apresentou resultado significativo entre as proporções (p-valor<0,01). Portanto, verificou-se que, com o uso do comitê de redes obtém-se uma PCP superior à proporção de acertos obtida pela equação de predição EqPred (Tabela 3).

Na Tabela 4, tem-se os resultados estatísticos associados aos modelos selecionados neste trabalho para a predição da EMAn de alimentos para aves de corte.

Tabela 2 Valores observados, intervalos de predição (IP) e intervalos de credibilidade HPD referentes aos dados do conjunto de teste de modelos de predição da EMAn de alimentos para aves

Alimento	EMAn obs.	EqPred* (IP)	Comitê de redes* (HPD)
1	3747	[3639; 3793]	[3537; 3828]
2	2373	[2219; 2850]	[2337; 3057]
3	3699	[3691; 3748]	[3538; 3803]
4	3813	[3674; 3754]	[3538; 3832]
5	3598	[3463; 3585]	[3184; 3707]
6	3529	[3452; 3563]	[3213; 3721]
7	3862	[3572; 3749]	[3204; 3818]
8	2682	[2679; 3628]	[2075; 3677]
9	3624	[3719; 4314]	[3212; 3984]
10	1941	[2081; 2403]	[1722; 2791]
11	3676	[3596; 3889]	[3213; 3840]
12	2326	[2007; 3064]	[2207; 3112]
13	2355	[1971; 2902]	[2263; 3005]
14	2396	[1722; 3116]	[2184; 3048]
15	2478	[1922; 2984]	[2254; 3047]
16	3159	[2602; 3466]	[2349; 3976]
17	3779	[3225; 4171]	[3073; 3906]
18	2809	[2040; 3161]	[2254; 3410]
19	3772	[3437; 4684]	[3204; 4407]
20	3934	[3536; 4436]	[2851; 4212]
21	3591	[3774; 3857]	[3538; 3831]
22	3353	[3501; 3634]	[3179; 3709]
23	2248	[1493; 3212]	[1668; 3604]
24	2387	[2302; 3051]	[2203; 3277]
25	3700	[3597; 4419]	[2851; 4238]
26	3971	[3513; 4170]	[3254; 3969]
27	3288	[3055; 4325]	[2882; 3969]
28	3568	[3774; 3857]	[3538; 3831]
29	3297	[3501; 3634]	[3179; 3709]
30	2151	[1493; 3212]	[1668; 3604]
31	2314	[2302; 3051]	[2203; 3277]
32	3941	[3597; 4419]	[2851; 4238]
33	3818	[3513; 4170]	[3254; 3969]
34	3173	[3055; 4325]	[2882; 3969]
35	3581	[3774; 3857]	[3538; 3831]
36	3436	[3501; 3634]	[3179; 3709]
37	2173	[1493; 3212]	[1668; 3604]
38	2339	[2302; 3051]	[2203; 3277]
39	3954	[3597; 4419]	[2851; 4238]
40	3793	[3513; 4170]	[3254; 3969]
41	3330	[3055; 4325]	[2882; 3969]
42	3548	[3774; 3857]	[3538; 3831]
43	3363	[3501; 3634]	[3179; 3709]
44	2162	[1493; 3212]	[1668; 3604]
45	2309	[2302; 3051]	[2203; 3277]
46	4022	[3597; 4419]	[2851; 4238]
47	3890	[3513; 4170]	[3254; 3969]
48	3267	[3055; 4325]	[2882; 3969]
Amplitude Média		723,741	934,521

Intervalo em negrito representa que não há nenhuma evidência de que o valor observado difere do valor predito pelo modelo, para o respectivo

alimento. *O comitê de redes foi obtido em Mariano et al. (2013a). A EqPred foi obtida utilizando-se a metodologia proposta neste trabalho.

Tabela 3 Porcentagem de acerto de predição (PCP) de modelos para a predição da EMAn de alimentos para aves de corte, considerando dados de teste obtidos em ensaio metabólico

Modelos*	PCP** (%)
EqPred	72,92 a
Comitê de redes neurais	95,83 b

* O modelo de redes neurais foi obtido em Mariano et al. (2013a). A equação EqPred foi obtida utilizando-se a metodologia proposta neste trabalho.

** Porcentagens seguidas de letras iguais na coluna, não diferem entre si pelo teste assintótico de χ^2 de Pearson para comparação de proporções, com 5% de significância.

Tabela 4 Estatísticas usadas para verificar a adequabilidade do ajuste e a acurácia dos modelos selecionados

Estatísticas*	Modelos**	
	EqPred	Comitê de redes
R ²	0,87	0,89
m.s.e.	54629,21	45285,43
MAD	190,83	175,66
MAPE (%)	6,52	5,97
Bias	-146,67	-86,80

* R² - coeficiente de determinação; MAD - desvio médio absoluto; MAPE - erro médio percentual absoluto; m.s.e. - erro quadrático médio. ** O modelo de redes neurais foi obtido em Mariano et al. (2013a). A equação EqPred foi obtida utilizando-se a metodologia proposta neste trabalho.

O modelo de comitê de redes neurais possui os menores valores de MAD, MAPE e *m.s.e.*, valor de R² mais próximo de um e Bias mais próximo de zero, quando comparado à EqPred (Tabela 4). Estes resultados, obtidos pelos critérios de adequabilidade de modelos, demonstraram que o comitê de redes fornece predições mais acuradas frente ao modelo de RLMP ajustado, o que condiz com o fato de que o modelo de comitê garante uma maior PCP. Além do mais, a tendência de variáveis independentes serem fortemente correlacionadas umas com as outras e a existência de até interações quádruplas significativas, faz com que

a aplicação prática da EqPred seja inviável para os nutricionistas da área animal. Em contrapartida, no comitê é possível utilizar muitas variáveis que interferem na variabilidade da energia metabolizável, o que nem sempre é viável em um modelo de regressão múltipla.

Portanto, o comitê composto de $N=1000$ redes, considerando como melhor estimador a moda empírica das predições, possuiu uma boa generalização e será disponibilizado para o uso com novos dados de entrada. Para isso, pretende-se criar uma ferramenta de rápida e fácil utilização para os nutricionistas da área animal (MARIANO et al., 2013b).

O gráfico da Figura 1 mostra o desempenho do comitê de redes, comparando o valor predito com o valor real.

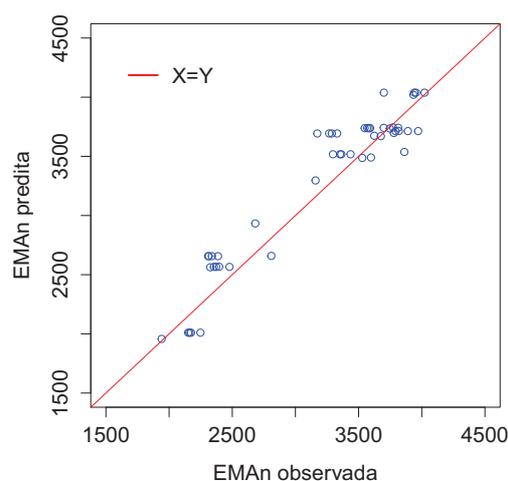


Figura 1 Valores de EMAn preditos vs. observados obtidos nos dados de teste para alimentos concentrados para frangos de corte, via comitê de redes neurais

Os valores preditos pelo comitê de redes tenderam a seguir os mesmos padrões que os dados observados (Figura 1), pois pontos próximos à linha de iden-

tidade indicam que os valores preditos são mais semelhantes aos observados. Ou seja, foi possível verificar o bom ajuste do modelo para a predição da EMAn de alimentos concentrados energéticos e proteicos, comumente utilizados na formulação de rações para aves de corte.

Os resultados obtidos neste artigo reforçam as conclusões de estudos anteriores, de que o uso de modelos relacionados à ANN garante predições mais acuradas do que as obtidas por métodos estatísticos clássicos, na área de nutrição e de produção animal. Recentemente, Perai et al. (2010) constataram que o modelo de ANN produziu predições mais acuradas de valores de energia metabolizável verdadeira (EMVn) de amostras de farinha de carne e ossos, quando comparado com modelos de mínimos quadrados parciais e regressão linear múltipla. Wang, Chen e Roan (2012) mostraram que o método de ANN é mais acurado do que os modelos tradicionais de regressão para a predição de produção de ovos em Taiwan. Ahmadi e Golian (2010) verificaram que o modelo de rede neural de base radial (RBFN) apresentou predições mais acuradas do crescimento de frangos de corte do que às obtidas por modelos de regressão múltipla, sendo esses frangos alimentados com dietas variando o percentual de energia metabolizável fornecido por proteína, gordura e carboidrato.

Em estudos futuros, pretende-se utilizar algum método que possibilite avaliar a significância das variáveis de entrada a serem consideradas nas redes especialistas que compõem o comitê, como *bayesian networks*. O objetivo será o de aprimorar a acurácia do modelo de comitê de redes na predição da EMAn para alimentos concentrados utilizados na formulação de rações avícolas. Além disso, novos modelos de regressão deverão ser ajustados e uma análise mais detalhada sobre intervalos de predição deverá ser considerada.

4 CONCLUSÃO

A utilização do comitê de redes neurais apresentou resultados superiores aos obtidos pela equação de predição ajustada neste estudo. Assim, o uso do comitê com N=1000 redes e utilizando-se a moda da distribuição empírica das predições na combinação dos resultados, garantiu uma maior acurácia nas predições do conjunto de teste. Portanto, a utilização desse modelo de comitê de redes é recomendado para a predição de valores acurados de EMAn de alimentos concentrados energéticos e proteicos, utilizados comumente em dietas de frangos de corte.

AGRADECIMENTOS

Os autores agradecem o apoio financeiro da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES).

REFERÊNCIAS

- AHMADI, H.; GOLIAN, A. Growth analysis of chickens fed diets varying in the percentage of metabolizable energy provided by protein, fat, and carbohydrate through artificial neural network. **Poultry Science**, Champaign, v. 89, n. 1, p. 173-179, Jan. 2010.
- ALVARENGA, R. R. et al. Energetic values of feedstuffs for broilers determined with in vivo assays and prediction equations. **Animal Feed Science and Technology**, Amsterdam, v. 168, p. 257-266, Sept. 2011.
- ANYAECHE, C. O.; IGHRAVWE, D. E. Predicting performance measures using linear regression and neural network: a comparison. **African Journal of Engineering Research**, Surulere, v. 1, n. 3, p. 84-89, July 2013.
- BIASE, N. G.; FERREIRA, D. F. Comparações múltiplas e testes simultâneos para parâmetros binomiais de k populações independentes. **Revista Brasileira de Biometria**, São Paulo, v. 27, n. 3, p. 301-323, jul./set. 2009.

BONAT, W. H. et al. **Métodos computacionais para inferência estatística**. João Pessoa: Associação Brasileira de Estatística, 2012. 289 p.

FERREIRA, D. F. **Estatística básica**. Lavras: UFLA, 2005. 664 p.

HAYKIN, S. **Neural networks: a comprehensive foundation**. 3rd ed. Upper Saddle River: Prentice Hall, 2007.

MARIANO, F. C. M. Q. et al. Equações de predição de valores energéticos de alimentos obtidas utilizando meta-análise e componentes principais. **Ciência Rural**, Santa Maria, v. 42, n. 9, p. 1634-1640, set. 2012.

MARIANO, F. C. M. Q. et al. Neural network committee to predict the AMEn of poultry feedstuffs. **Neural Computing and Applications**, London, 2013a. In press.

MARIANO, F. C. M. Q. et al. Prediction of the energy values of feedstuffs for broilers using meta-analysis and neural networks. **Animal**, Cambridge, v. 7, n. 9, p. 1440-1445, Sept. 2013b.

NASCIMENTO, G. A. J. et al. Equações de predição para estimar os valores energéticos de alimentos concentrados de origem vegetal para aves utilizando a metanálise. **Revista Brasileira de Zootecnia**, Viçosa, MG, v. 38, n. 7, p. 1265-1271, dez. 2009.

PERAI, A. H. et al. A comparison of artificial neural networks with other statistical approaches for the prediction of true metabolizable energy of meat and bone meal. **Poultry Science**, Champaign, v. 89, p. 1562-1568, July 2010.

R DEVELOPMENT CORE TEAM. **R: a language and environment for statistical computing**. Vienna: R Foundation for Statistical Computing, 2013. Software.

RODRIGUES, P. B. et al. Valores energéticos da soja e subprodutos da soja, determinados com frangos de corte e galos adultos. **Revista Brasileira de Zootecnia**, Viçosa, MG, v. 31, n. 4, p. 1771-1782, jul./ago. 2002.

SHI, H. et al. Comparison of artificial neural network and logistic regression models for predicting in-hospital mortality after primary liver cancer surgery. **PLoS ONE**, San Francisco, v. 7, n. 4, Apr. 2012. Disponível em: <<http://www.plosone.org/article/info%3Adoi%2F10.1371%2Fjournal.pone.0035781>>. Acesso em: 10 dez. 2013.

WANG, B. Y.; CHEN, S. A.; ROAN, S. W. Comparison of regression and artificial neural network on egg production. **Journal of Animal and Veterinary Advances**, Kuala Lumpur, v. 11, n. 14, p. 2503-2508, 2012.

CONSIDERAÇÕES GERAIS

A utilização de redes neurais para realizar a predição de valores de EMAn para alimentos utilizados na formulação de rações para aves de corte é um método promissor de análise e auxílio na tomada de decisões nesta área abordada.

O modelo MLP mais acurado possui sete entradas, cinco neurônios na primeira camada intermediária, três neurônios na segunda camada intermediária e um neurônio como saída. Com base nesses resultados, construiu-se uma macro para a predição da EMAn, sendo acessível e de fácil manuseio para os nutricionistas da área animal.

Pelo uso de intervalos HPD, no conjunto de teste, foi possível considerar incerteza nas predições do comitê de redes. Além disso, o uso do comitê, em que N=1000 redes foram combinadas pelo estimador da moda da distribuição empírica das predições, garantiu uma maior porcentagem de acerto das predições do conjunto de teste (PCP=95,83%), o que é bastante satisfatório.

A utilização deste comitê de redes neurais apresentou valor de PCP superior ao obtido pelo modelo de regressão linear múltipla ponderada, considerando os efeitos principais e de interações entre as variáveis explicativas do modelo. Além disso, pela análise das medidas de adequabilidade, o comitê de redes garantiu predições mais acuradas. Portanto, com base nos resultados deste comitê, pretende-se criar uma ferramenta de rápida e fácil utilização para os nutricionistas da área animal obterem a predição acurada da EMAn de alimentos concentrados energéticos e proteicos, utilizados comumente em dietas de frangos de corte.

O modelo de comitê apresentado nesta tese é passível de atualizações, pretendendo-se, em estudos futuros, utilizar algum método que possibilite avaliar a significância das variáveis de entrada a serem consideradas nas ANN que

compõem o comitê. Também planeja-se atualizar os dados com resultados experimentais internacionais. O objetivo será o de aprimorar a acurácia do comitê de redes na predição da EMAn, a fim de obter um modelo que seja o mais acurado possível.