

APLICAÇÃO DE REDES NEURAIS ARTIFICIAIS COMO TESTE DE DETECÇÃO DE FRAUDE DE LEITE POR ADIÇÃO DE SORO DE QUEIJO

Applying artificial neural networks as a test to detect milk fraud by whey addition

Gerson de Freitas Silva Valente^{1}, Daiana Cristina Guimarães¹,
Ana Laís Andrade Gaspardi¹, Lara de Andrade Oliveira¹*

RESUMO

Esse trabalho foi realizado com o objetivo de empregar Redes Neurais Artificiais para classificar amostras de leite, a partir de análises de rotina de um laticínio, em amostras de leite adulteradas ou não, quanto à adição de soro de queijo. As amostras foram preparadas através da mistura do leite com diferentes concentrações de soro (0, 1, 5, 10 e 20%) sendo analisadas quanto a temperatura, teor de gordura, extrato seco desengordurado, densidade, proteína, sais minerais, ponto de congelamento, condutividade e pH, totalizando 167 ensaios. Desses 167 ensaios, 101 foram usados para treinamento da rede, 33 para validação e outros 33 para testar a rede neural artificial. A melhor rede de classificação foi uma rede neural de função de base radial. Para obter os centros da rede foi usado o algoritmo *k*-means, para definir a largura dos campos receptivos o *k*-nearest e os pesos da camada de saída foram definidos usando o método da pseudo-inversa. A rede com melhor resultado apresentou 10 neurônios na camada de entrada, 40 neurônios na camada oculta e dois na camada de saída, sendo possível obter mais de 95% de acertos na classificação. A metodologia de classificação por Redes Neurais Artificiais apresenta um grande potencial de aplicação na interpretação de dados de análises de rotina nos laticínios para classificação do leite em adulterado por soro de queijo e, posteriormente, confirmação do resultado por metodologia oficial.

Palavras-chave: classificação; análise físico-química; RBF.

1 Instituto Federal de Educação, Ciência e Tecnologia, Sul de Minas Gerais (IFSULDEMINAS), Câmpus Inconfidentes, Praça Tiradentes, 416, Centro, 37576-000, Inconfidentes, MG, Brasil. E-mail: valentegerson@hotmail.com

* Autor para correspondência.

Recebido / Received: 06/08/2014

Aprovado / Approved: 11/11/2014

ABSTRACT

This study aimed to employ Artificial Neural Networks to classify milk samples from routine analysis of a dairy company in order to identify adulteration by whey addition. The samples were prepared by mixing the milk with different whey concentrations (0, 1, 5, 10, and 20%), which were then analyzed for temperature, fat content, solids-non-fat, bulk density, protein, lactose, minerals, freezing point, conductivity, and pH, for a total of 167 assays. Out of these, 101 were used to train the network, 33 for validation, and 33 to test the artificial neural network. The best classification was obtained using a radial basis function neural network. *k*-means algorithm was used to obtain the network center, *k*-nearest was used to define the receptive fields, and the pseudo-inverse method was used to define the weights of the output layer. The best result was found with a network with 10 neurons in the input layer, 40 neurons in the hidden layer, and two neurons in the output layer, achieving over 95% accuracy in classification. The classification methodology using artificial neural networks has strong potential to be applied in interpreting data from routine analysis in dairy companies in order to classify milk adulterated with whey and, later, confirm the result using official methodologies.

Keywords: classification; physicochemical analysis; RBF.

INTRODUÇÃO

A adulteração de leite por meio da adição intencional de soro de leite é um problema que vem ocorrendo devido ao baixo valor comercial do soro, seu reduzido aproveitamento em derivados e subprodutos lácteos e do alto custo para o tratamento do efluente antes do lançamento em corpos receptores (MAGALHÃES, 2008).

Segundo Richards (1997), a quantidade do soro de leite gerada no processo de fabricação de queijo é função do tipo de queijo produzido e das técnicas de fabricação empregadas, mas em média, para a fabricação de um quilo de queijo são gerados de 9 a 12 litros de soro de queijo.

A adição fraudulenta de soro de queijo ao leite é normalmente detectada e quantificada pela determinação do caseinomacropéptido (CMP), um fragmento hidrofílico da *k*-caseína, liberado pela ação da quimosina durante a coagulação enzimática do leite que permanece solúvel no soro, e que deve estar ausente no leite. Atualmente existem vários métodos para

a detecção desta fraude, sendo que no Brasil as atenções têm-se voltado para a fraude por adição de soro ao leite *in natura*, devido a sua maior participação no mercado nacional (PEREIRA et al., 2001; SILVA et al., 2009a).

A CMP é geralmente quantificada através de métodos cromatográficos (VELOSO et al., 2002), eletroforéticos (SOUZA et al., 2000) e espectrofotométricos (FUKUDA et al., 1996). No entanto, essas metodologias apresentam custo elevado e requerem equipamentos específicos e necessitam de mão de obra capacitada para realização das análises.

As análises de rotina normalmente não detectam a presença de soro, a não ser quando adicionado em grandes quantidades (MAGALHÃES, 2008). Alguns pesquisadores, Soares et al. (2005 e 2006), Silva et al., (2009a e 2009b), Silva et al. (2010), têm avaliado o potencial de aplicação da estatística multivariada (MANOVA), especificamente da análise discriminante, na avaliação e interpretação de dados de rotina na indústria de laticínios.

Uma alternativa para classificação de amostras são as Redes Neurais Artificiais. As Redes Neurais Artificiais podem funcionar como modelos preditivos que descrevem a relação funcional entre as variáveis de entrada e variáveis de saída de um sistema. As Redes Neurais Artificiais (RNA) têm várias vantagens sobre os modelos fenomenológicos tradicionais ou modelos empíricos. RNA desenvolvem um mapeamento das variáveis de entrada e saída, que podem ser usados para prever parâmetros de saída do sistema (SINGH et al., 2009).

As RNA procuram desenvolver modelos computacionais baseados na capacidade do cérebro humano. As principais propriedades estão relacionadas à capacidade de aprender por exemplos, de interpolar com base em padrões fornecidos e de selecionar propriedades específicas dentro do universo amostral (ABER et al., 2009; SINGH et al., 2009; ELMOLLA et al., 2010; KHATAEE et al., 2010).

A unidade básica para o processamento de informações é o neurônio artificial que pode receber uma ou mais entradas transformando-as em saídas. Cada entrada tem um peso associado que determina sua intensidade de influência no dado de saída (ABER et al., 2009; ELMOLLA et al., 2010; KHATAEE et al., 2010).

Na definição da arquitetura de uma RNA, consideram-se o número de camadas, o número de neurônios e a conexão entre os neurônios (DANESHVAR et al., 2006; ABER et al., 2009; SINGH et al., 2009; ELMOLLA et al., 2010).

Esse trabalho tem como objetivo empregar Redes Neurais Artificiais para classificar amostras de leite em leite adulterado ou normal quanto à adição de soro de queijo a partir de análises de rotina de um laticínio.

MATERIAL E MÉTODOS

O trabalho foi desenvolvido na fazenda escola do Instituto Federal do Sul de Minas

Gerais – Câmpus Inconfidentes. Os ensaios laboratoriais foram realizados no setor de laticínios.

Amostras de leite em diferentes dias de produção (95 amostras sem adulteração) foram usadas para configuração da RNA, sendo que 18 amostras foram adulteradas adicionando 1, 5, 10 e 20% de soro de queijo num total de 72 amostras como forma de detectar as possíveis fraudes nas amostras. As análises de rotina de um laticínio do leite e das amostras adulteradas foram realizadas através de um analisador de leite por ultrassom (*MASTER COMPLETE-AKSO®*). As análises foram: teor de gordura, extrato seco desengordurado, densidade, proteína, lactose, sais minerais, ponto de congelamento, condutividade, pH e temperatura.

O soro de queijo foi proveniente da produção de queijo Minas Padrão e Muçarela. Os soros provenientes da produção desses queijos não apresentam diferenças significativas em suas propriedades (TEIXEIRA; FONSECA, 2008). O soro foi filtrado em desadorador para remoção de partículas grosseiras da massa do queijo.

Para o desenvolvimento da arquitetura da RNA, os 167 ensaios foram distribuídos aleatoriamente nos subgrupos: treinamento (60%), validação (20%) e teste (20%), de acordo com Valente et al. (2014). Para o subgrupo de treinamento foram 53 amostras de leite normal e 48 amostras adulteradas com soro de queijo. Para validação e teste foram 21 casos de leite normal e 12 de leite adicionado de soro distribuídos de forma aleatória.

Os valores das variáveis foram normalizadas entre 0 e 1 usando a Equação 1.

$$X_i = \frac{x - \min(x)}{\max(x) - \min(x)} \quad (1)$$

em que, x é valor observado, $\min(x)$ o menor valor para a variável e $\max(x)$ o maior valor da variável em estudo.

Para determinar a configuração da rede neural artificial foi usado o *software* SNN[®]. Várias redes foram testadas, redes de função radial e *multilayer perceptron* com uma camada oculta. Como critério de seleção para as redes foi usado o erro quadrático médio (EQM) para os dados de validação (CARVALHO et al., 2013). O número de neurônios nessas camadas foram definidos através de tentativas e erro.

As variáveis de entrada foram as variáveis obtidas em análises de rotina: temperatura, pH, teor de gordura, extrato seco desengordurado, proteínas, ponto de congelamento, condutividade, lactose e densidade das amostras. Duas configurações foram testadas para a camada de saída: uma com um neurônio em que leite normal era 0 (zero) e adulterado era 1 (um); outra com dois neurônios, leite normal (0, 1) e leite adulterado (1, 0).

Para obter os centros da rede (RBF) foi usado o algoritmo *k-means*, para definir a largura dos campos receptivos o *k-nearest* e os pesos da camada de saída foram definidos usando o método da pseudo-inversa. Para as redes *Multilayer perceptron* (MLP), o treinamento foi realizado usando o algoritmo retropropagação (*backpropagation*).

As dez melhores RNA foram testadas

para classificação do leite em normal ou adulterado. A melhor rede foi aquela que apresentou menor erro de classificação.

RESULTADOS E DISCUSSÃO

As amostras de leite e leite adulterado com soro de queijo apresentaram características que estão na Tabela 1.

Como critério de seleção para as redes foi usado o erro quadrático médio (EQM) para os dados de validação (CARVALHO et al., 2013). De 28.017 redes testadas, entre *Multilayer perceptron* e RBF, as melhores estão na Tabela 2. As redes com dois neurônios na camada de saída apresentaram melhores resultados.

As RNA apresentadas na Tabela 2 foram testadas para classificação das amostras. A melhor rede foi aquela que apresentou menor número de erros de classificação, menor diferença entre os erros de classificação para treinamento, validação e teste, e como último critério, o menor número de neurônios na camada oculta.

A melhor rede de classificação foi uma rede neural de função de base radial. A rede com melhor resultado apresentou 10 neurônios na camada de entrada, 40 neurônios

Tabela 1 – Características das amostras usadas para configurar a RNA

	Mínimo	Máximo	Média	Desvio-padrão
Temperatura (°C)	7,9	27,6	14,2	3,61
Gordura (%)	1,65	6,06	3,0	0,70
ESD (%)	5,4	8,85	8,2	0,78
Densidade (kg/m ³)	1019,8	1031,9	1030,0	2,78
Proteína (%)	1,95	3,16	3,0	0,28
Lactose (%)	2,97	4,89	4,5	0,43
Sais minerais (%)	0,45	0,77	0,7	0,07
Ponto de congelamento (°C)	-0,586	-0,327	-0,521	0,06
Condutividade (mS/cm)	4,5	5,2	4,9	0,18
pH	6,4	7,0	6,6	0,13

na camada oculta e 2 na camada de saída (Figura 1).

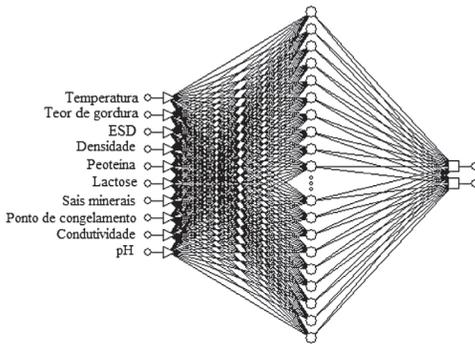


Figura 1 – Estrutura da RBF usada na detecção de adição de soro de queijo no leite

Com a rede neural artificial da Figura 1 foi possível classificar as amostras em leite normal ou adulterado pela adição de soro de queijo. Os resultados estão apresentados na Tabela 3.

Dos 167 casos ocorreu erro de classificação em oito, ou seja, erro de 4,8 %. Para o leite que não foi adulterado (0%) não ocorreu erro de classificação. Já para amostras adulteradas houve classificação errônea, mas numa porcentagem considerada baixa quanto comparada a outros trabalhos que utilizaram a metodologia estatística multivariada para classificação.

Esses resultados mostram que as Redes Neurais Artificiais apresentam grande

Tabela 2 – Redes Neurais Artificiais testadas com menor EQM para os dados de validação

Tipo	Neurônios na camada de entrada	Neurônios na camada oculta	EQM validação	Treinamento
RBF	10	40	0,274	KM,KN,PI
RBF	10	51	0,371	KM,KN,PI
RBF	10	50	0,387	KM,KN,PI
RBF	10	49	0,388	KM,KN,PI
RBF	10	47	0,392	KM,KN,PI
RBF	10	43	0,396	KM,KN,PI
RBF	10	21	0,418	KM,KN,PI
RBF	10	21	0,419	KM,KN,PI
RBF	10	19	0,420	KM,KN,PI
RBF	10	21	0,421	KM,KN,PI
RBF	7	21	0,421	KM,KN,PI
RBF	10	21	0,424	KM,KN,PI
RBF	5	21	0,425	KM,KN,PI
RBF	10	21	0,427	KM,KN,PI
RBF	10	21	0,427	KM,KN,PI
RBF	10	20	0,427	KM,KN,PI
MLP	10	20	0,591	<i>Backpropagation</i>
MLP	10	31	0,609	<i>Backpropagation</i>
MLP	1	31	0,667	<i>Backpropagation</i>
MLP	1	13	0,766	<i>Backpropagation</i>

KM é *k-means*, KN *k-nearest*, PI pseudo-inversa.

Tabela 3 – Erros de classificação pela RNA selecionada

	Erros	%	Casos	Erros nas concentrações				
				0%	1%	5%	10%	20%
Treinamento	4	4,0	101	0	1	0	1	2
Validação	2	6,1	33	0	2	0	0	0
Teste	2	6,1	33	0	0	1	0	1
Total	8	4,8	167	0	3	1	1	3

potencial de uso como ferramenta de avaliação de possíveis adulterações de leite com soro de queijo. Essa técnica apresenta uma grande possibilidade de aplicação reduzindo o número de amostras que iriam ser analisadas pela metodologia oficial.

Aplicando a metodologia da estatística multivariada (MANOVA), Silva et al. (2009b) obtiveram uma margem de erro de 36,0%. Já Silva et al. (2010), trabalhando com amostras coletadas em Viçosa, MG, no período de junho de 2009 a março de 2010, obtiveram uma margem de erro de 30,3%. Os resultados desse artigo demonstram erro de classificação menor que 5% aplicando a rede neural artificial.

Uma grande vantagem da rede neural artificial é que pode ser aprimorada, quanto maior o número de dados disponíveis melhor será a predição da rede. Novas variáveis podem ser incorporadas na rede, como por exemplo, o produtor e estação do ano.

Maia et al (2010) realizaram a avaliação das características sensoriais e físico-químicas nas fraudes por adição de água, soro de queijo, soro fisiológico e soro glicosado. Concluíram que na análise sensorial, apenas foi detectada diferença a partir de 25% de água adicionada, 45% de soro de queijo, 20% de soro fisiológico e 30% de soro glicosado. As adições reduziram os teores de gordura, lactose, proteínas e minerais. A crioscopia foi eficaz para detectar aguagem, porém a adição de solutos dificultou a detecção. A pesquisa mostrou a influência das

substâncias adicionadas nas características físico-químicas do leite. Essa variação na composição pode ser a chave para a detecção das fraudes usando Redes Neurais Artificiais.

CONCLUSÕES

A utilização de rede neurais artificiais permite a detecção de fraude por adição de soro de queijo ao leite a partir dos resultados obtidos na análise de rotina.

Os resultados obtidos demonstram um grande potencial de aplicação das Redes Neurais Artificiais para seleção das possíveis amostras adulteradas com soro de queijo.

AGRADECIMENTOS

Ao CNPq.

REFERÊNCIAS

- ABER, S.; AMANI-GHADIM, A. R.; MIRZAJANI, V. Removal of Cr (VI) from polluted solutions by electrocoagulation: Modeling of experimental results using artificial neural network. **Journal of Hazardous Materials**, v. 171, n. 1-3, p. 484-490, 2009.
- CARVALHO, N. B. et al. Artificial neural networks (ANN): prediction of sensory measurements from instrumental data. **Food Science and Technology**, v. 33, n. 4, p. 722-729, 2013.

- DANESHVAR, N.; KHATAEE, A. R.; DJAFARZADEH, N. The use of artificial neural networks (ANN) for modeling of decolorization of textile dye solution containing C.I. Basic Yellow 28 by electro-coagulation process. **Journal of Hazardous Materials**, v. 137, n. 3, p. 1788-1795, 2006.
- ELMOLLA, E. S.; CHAUDHURI, M.; ELTOUKHY, M. M. The use of artificial neural network (ANN) for modeling of COD removal from antibiotic aqueous solution by the Fenton process. **Journal of Hazardous Materials**, v. 179, n. 1-3, p. 127-134, 2010.
- FUKUDA, S. P.; ROIG, S. M.; PRATA, L. F. Aplicação do método da ninidrina ácida como teste de "screening" de plataforma para detecção da adição de soro ao leite. **Ciência e Tecnologia de Alimentos**, v. 16, n. 1, p. 52-56, 1996.
- KHATAEE, A. R.; ZAREI, M.; POURHASSAN, M. Bioremediation of malachite green from contaminated water by three microalgae: Neural network modeling. **Clean- Soil, Air, Water**, v. 38, n. 1, p. 96-103, 2010.
- MAGALHÃES, M. A. **Determinação de fraude de leite com soro de leite pela análise de CMP e PSEUDO-CMP por cromatografia líquida de alta eficiência em fase reversa com detecção por espectrometria de massa**. 2008. 57 f. Dissertação (Mestrado em Ciência e Tecnologia de Alimentos) – Universidade Federal de Viçosa, Viçosa, 2008.
- MAIA, R. G.; CALIL, C.; COSTA, A. Características físico-químicas e análise sensorial do leite pasteurizado adicionado de água, soro de queijo, soro fisiológico e soro glicosado. **Revista do Instituto de Laticínios Cândido Tostes**, v. 65, n. 376, p. 18-25, 2010.
- PEREIRA, D. B. C. et al. **Físico Química do Leite e Derivados: Métodos analíticos**, 2. ed. Juiz de Fora: EPAMIG, 2001. 234p.
- RICHARDS, N. S. P. S. Uso racional de soro lácteo. **Revista Indústria de Laticínios**, v. 2, n. 9, p. 67-69, 1997.
- SILVA, R. A. G. et al. Características físico-químicas de leite cru fraudado com diferentes concentrações de soro de leite. In: ENCONTRO NACIONAL DE ANALISTAS DE ALIMENTOS (ENAAL), 16., 2009. **Anais...** Belo Horizonte, 2009a. 1 CD-ROM.
- SILVA, R. A. G. et al. Detecção da adição de soro ao leite cru através da análise estatística multivariada. In: ENCONTRO NACIONAL DE ANALISTAS DE ALIMENTOS (ENAAL), 16., 2009. **Anais...** Belo Horizonte, 2009b. 1 CD-ROM.
- SILVA, T. D. da et al. Aplicação da análise multivariada para obtenção de uma função discriminante para detecção de adição de soro de leite em leite. In: CONGRESSO NACIONAL DE LATICÍNIOS, 27., 2010. **Anais...** Juiz de Fora: Templo, 2010. 1 CD-ROM.
- SINGH, K. P. et al. Artificial neural network modeling of the river water quality: A case study. **Ecological Modelling**, v. 220, v. 6, p. 888-895, 2009.
- SOARES, E. C. et al. Análise estatística multivariada como teste para detecção de adição de soro ao leite pasteurizado. **Revista do Instituto de Laticínios Cândido Tostes**, n. 351, v. 61, p. 37-42, 2006.
- SOARES, E. C., RAMOS, A. L. S., RAMOS, E. M. Detecção de fraude de leite através de análise multivariada. In: SIMPÓSIO

LATINO AMERICANO DE CIÊNCIA DOS ALIMENTOS (SLACA), 6., 2005. **Anais...** Campinas: SBCTA, 2005. 1 CD-ROM.

SOUZA, E. M. T. de et al. Electrophoretic analysis to detect and quantify additional whey in milk and dairy beverages. **Ciência e Tecnologia de Alimentos**, v. 20, n. 3, p. 314-317, 2000.

TEIXEIRA, L. V; FONSECA, L. M. Perfil físico-químico do soro de queijos mozzarella e minas-padrão produzidos em várias regiões do estado de Minas Gerais. **Arquivo Brasileiro de Medicina Veterinária e**

Zootecnia, v. 60, n. 1, p. 243-250, 2008.

VALENTE, G. F. S. et al. Artificial neural network prediction of chemical oxygen demand in dairy industry effluent treated by electrocoagulation. **Separation and Purification Technology**, v. 132, p. 627-633, 2014. Disponível em: <<http://linkinghub.elsevier.com/retrieve/pii/S1383586614003670>>. Acesso em: 3 set. 2014.

VELOSO, A. C. A. et al. Detecção de adulterações em produtos alimentares contendo leite e/ou proteínas lácteas. **Química Nova**, v. 25, n. 4, p. 609-615, 2002.