

SISTEMAS INTELIGENTES NA CLASSIFICAÇÃO DA CARNE SUÍNA

LETHÍCIA OLÍMPIO BUENO^{1*}, ALFREDO BENETI DIANI², ALCINÉIA DE LEMOS SOUZA RAMOS³, EDUARDO MENDES RAMOS⁴, DANTON DIEGO FERREIRA⁵.

¹MSc. em Ciência dos Alimentos, Doutoranda em Ciência dos Alimentos (bolsista Fapemig), UFLA, Lavras-MG, lethicia.olimpio@outlook.com, *autor correspondente;

²Bacharel em Sistemas de Informação, Mestrando em Engenharia de Sistemas e Automação, UFLA, Lavras-MG, alfredodiani@outlook.com;

³Dr. em Ciência e Tecnologia de Alimentos, Profa. Associada, UFLA, Lavras-MG, alcineia@ufla.br;

⁴Dr. em Ciência e Tecnologia de Alimentos, Prof. Titular, UFLA, Lavras-MG, emramos@ufla.br;

⁵Dr. em Engenharia Elétrica, Prof. Associado, UFLA, Lavras-MG, danton@ufla.br.

Apresentado no
Congresso Técnico Científico da Engenharia e da Agronomia – CONTECC
7 a 10 de outubro de 2024

RESUMO: A qualidade de carne suína e a adequada classificação da mesma são de grande interesse para a indústria e cientistas. Este estudo visou aprimorar a classificação da qualidade da carne suína por meio de abordagens computacionais. Setenta suínos foram analisados, e um conjunto de dados inicialmente desbalanceado foi ajustado utilizando a técnica SMOTE (*Synthetic Minority Oversampling Technique*). A razão de discriminação de Fisher destacou que as características cruciais para distinguir classes de qualidade são L* e perda por gotejamento (PPG-DL). O método *K-means*, aplicado a um conjunto de dados sintético (BDs), revelou clara separação entre as classes. O estudo destaca o potencial dos sistemas inteligentes na avaliação da qualidade da carne suína, ressaltando a importância de características específicas e a eficácia de abordagens computacionais.

PALAVRAS-CHAVE: Qualidade de carne suína, PSE, Razão de Discriminação de Fisher, Reconhecimento de padrões, *K-means*.

INTELLIGENT SYSTEMS IN PORK MEAT CLASSIFICATION

ABSTRACT: The quality of pork and its proper classification are of great interest to both the industry and scientists. This study aimed to improve the classification of pork quality through computational approaches. Seventy pigs were analyzed, and an initially unbalanced dataset was adjusted using the SMOTE technique (*Synthetic Minority Oversampling Technique*). Fisher Discriminant Ratio highlighted that the crucial characteristics for distinguishing quality classes are L* and *drip loss* (PPG-DL). The *K-means* method, applied to a synthetic dataset (BDs), revealed clear separation between the classes. The study emphasizes the potential of intelligent systems in assessing pork quality, highlighting the importance of specific characteristics and the effectiveness of computational approaches.

KEYWORDS: Pork quality, PSE, Fisher Discriminant, Pattern recognition, *K-means*.

INTRODUÇÃO

A qualidade da carne suína é um tema de interesse para cientistas e a indústria de carne, pois está associada a fatores genéticos, nutricionais, de estresse pré-abate e processamento pós-abate (Torres Filho *et al.*, 2018). A carne suína pode ser classificada em diferentes categorias de qualidade, dependendo de sua cor, textura e exsudação (Peres *et al.*, 2018), sendo que a carne de alta qualidade é rosa-avermelhada, firme e não exsudativa (RFN). Condições anômalas podem resultar em carne pálida, macia e exsudativa (PSE), carne rosa-avermelhada, macia e exsudativa (RSE) ou carne escura, firme e seca (DFD), que apresentam má aparência e são inadequadas para processamento (Warner *et al.*, 1997; Faucitano *et al.*, 2010; Van Laack and Kauffman, 1999). A condição PSE é um desafio significativo para a indústria de carne, pois causa perdas econômicas expressivas. Autores propuseram classificar a carne suína entre três e cinco classes, mas ainda não há consenso sobre os melhores

critérios de classificação (Barbin *et al.*, 2012; Tomovic *et al.*, 2014; Warner *et al.*, 1997; Joo *et al.*, 2000; Kauffman *et al.*, 1993; Faucitano *et al.*, 2010; Torres Filho *et al.*, 2018). As classificações de carne suína propostas pelos autores utilizam limites rígidos e lógica clássica para definir as classes de qualidade, de modo que uma amostra só pode ser classificada se atender a todos os critérios de uma determinada classe, caso contrário, será considerada inviável, o que aumenta a dificuldade de classificar amostras que não se encaixam nos intervalos esperados dos parâmetros (Faucitano *et al.*, 2010; Chen e Li, 2010; Peres *et al.*, 2018).

Diferentes parâmetros podem ser usados como instrumento para medir a similaridade entre grupos, e diferentes métodos de agrupamento podem ser usados para agrupar amostras semelhantes (Torres Filho *et al.*, 2018). Sistemas inteligentes de aprendizado não supervisionado, como o *k-means*, emergem como uma solução eficaz. O *k-means* é capaz de agrupar amostras de carne suína em clusters com base na similaridade dos seus atributos, sem a necessidade de rótulos pré-definidos. Isso permite uma maior flexibilidade na classificação, especialmente para amostras que não se enquadram perfeitamente nas classes rígidas. Estudos demonstram que o *k-means* pode identificar padrões ocultos e proporcionar uma categorização mais natural e precisa das amostras (Xu e Tian, 2015; Jain, 2010). Este estudo objetiva desenvolver sistemas inteligentes capazes de selecionar características relevantes para a separação das classes e agrupá-los usando o *k-means*.

MATERIAL E MÉTODOS

Amostragem e Preparação das Amostras

Setenta suínos (Large White x Landrace) foram selecionados aleatoriamente durante sete dias (10 animais/dia), pesando 105 ± 10 kg, provenientes de diferentes produtores em um frigorífico inspecionado federalmente em Lavras, MG, Brasil. O músculo *Longissimus thoracis* foi retirado das carcaças 24 horas *post mortem*, embalado, armazenado em caixas térmicas e transportado para o LabCarnes na UFLA em 30 minutos. As amostras foram cortadas em bifes de aproximadamente 2,5 cm para análise de pH, cor instrumental e perda por gotejamento.

Análises de Qualidade de Carne

O pH final foi medido em triplicata utilizando um medidor portátil HI99.163 (Hanna Instruments Brasil, São Paulo, SP, Brasil) com eletrodo de vidro com ponta de lança. A cor instrumental foi avaliada 24 h *post mortem* com um colorímetro espectrofotômetro CM-700 (Konica Minolta, Inc., Tóquio, Japão) - abertura de 8 mm, SCE, iluminante/10°, obtendo-se as coordenadas L^* , a^* , b^* , C^* e h° , após 30 minutos de oxigenação das amostras. A perda por gotejamento foi medida 24 h *post mortem*, calculando a perda de peso de amostras padronizadas (cerca de 40 g) durante 48 h a 4°C, conforme Honikel *et al.* (1986), e expressa em porcentagem em relação ao peso inicial.

Banco de Dados

O banco de dados original foi composto pelos dados coletados dos 70 suínos, identificando as classes de qualidade de carne (DFD, PSE, RFN, RSE e PFN) segundo Torres Filho *et al.* (2018):

1. PSE: $L^* > 52$; *drip loss* > 6%
2. RSE: $L^* = 44$ à 52 ; *drip loss* > 6%
3. PFN: $L^* > 52$; *drip loss* = 2 - 6%
4. RFN: $L^* = 44$ - 52 ; *drip loss* = 2 - 6%
5. DFD: $L^* < 44$; *drip loss* < 2%

As features do banco de dados incluíram pH 24h, L^* , a^* , b^* , C^* , h e perda de peso por gotejamento (PPG-DL). As classes estavam desbalanceadas e com baixa quantidade de dados, consistindo em 25 animais (amostras) PSE, 20 RSE, 17 RFN, 6 PFN e 2 DFD.

Oversampling com SMOTE

Para lidar com o desbalanceamento das classes, utilizou-se a técnica SMOTE (Chawla *et al.*, 2002), gerando novas amostras das classes, totalizando 150 amostras por classe e um novo banco de dados (BDs) com 750 amostras.

Razão de Discriminação de Fisher

A Razão de Discriminação de Fisher (FDR) é uma técnica para redução da dimensionalidade em reconhecimento estatístico de padrões (Theodoridis & Koutroumbas, 2009). Para o uso do FDR problema foi tratado como uma classificação binária, confrontando cada classe contra todas as outras. O FDR é representado pela Equação:

$$(FDR) J = (\mu_1 - \mu_2)^2 \odot \frac{1}{\sigma_1^2 + \sigma_2^2}$$

Onde J é o número total de parâmetros, μ_1 e μ_2 , e σ_1^2 e σ_2^2 são, respectivamente, os vetores de média e variância dos dados da classe 1 e da classe 2, \odot refere-se ao produto de Hadamard.

K-means

O algoritmo *k-means*, proposto por MacQueen (1967), foi aplicado utilizando as features L* e PPG-DL, selecionadas pelo FDR. O *k-means* visa particionar N observações em k clusters. Cada observação é classificada no cluster que atingiu a menor distância euclidiana entre seu centroide e o vetor de observação (Silva *et al*, 2014). Foram formados cinco clusters, correspondendo às cinco classes conhecidas, com um máximo de 10.000 iterações e inicialização *k-means*.

RESULTADOS E DISCUSSÃO

Razão de Discriminação de Fisher

As *features* mais relevantes para todas as cinco classes de qualidade de carne suína foram L* e *drip-loss* (PPG-DL), embora também seja possível observar como se comporta as outras características. Esse resultado já era esperado, visto que já foi descrito por outros autores como características determinantes para classificação, porém, a abordagem pelo FDR consegue demonstrar o quanto elas são expressivas em cada uma das classes em relação as outras, demonstrando maior força do L* para discriminar a classe PSE e DFD das demais e o PPG-DL altamente determinante na separação entre os diferentes grupos. Os resultados gráficos da FDR são apresentados na Figura 1.

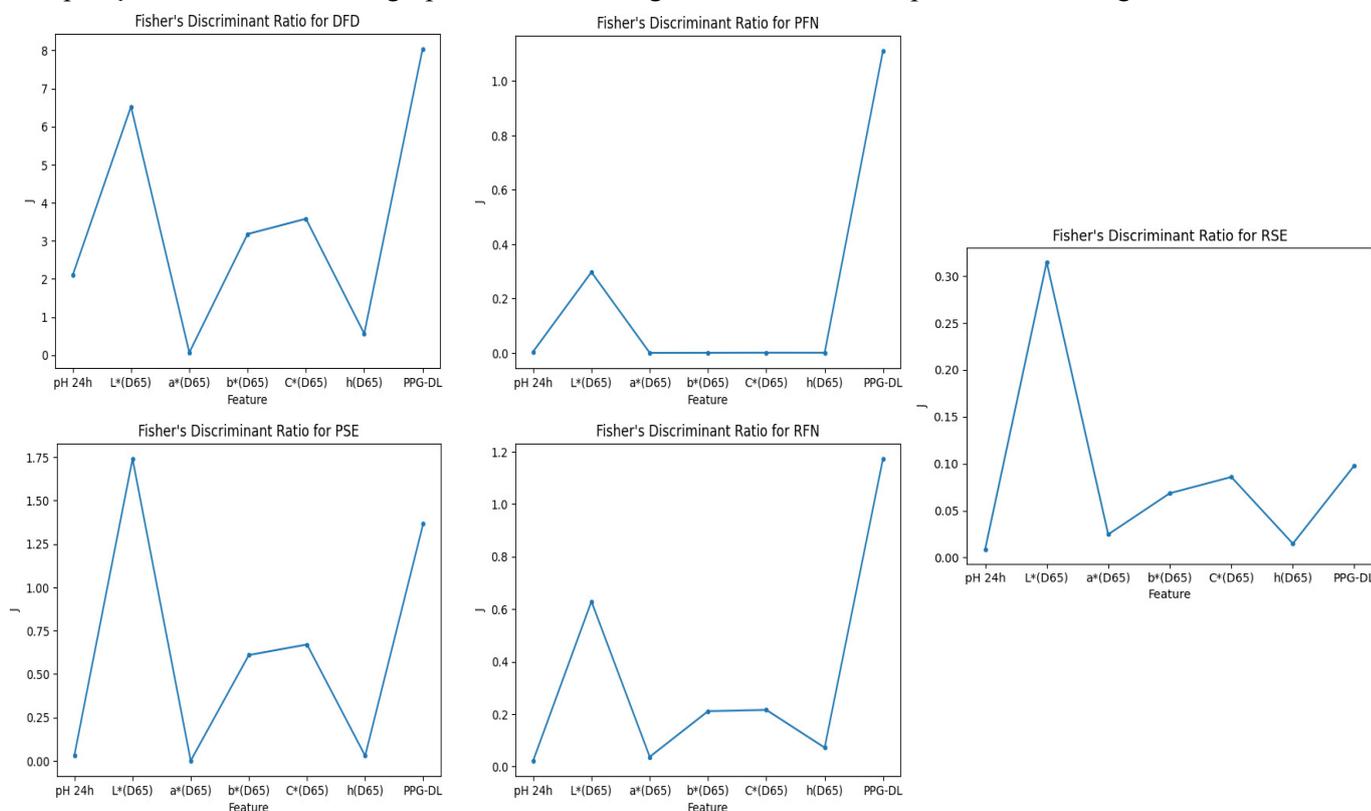


Figura 1: Resultados gráficos FDR

K-means

Foram feitas duas abordagens com o algoritmo *k-means*, primeiramente inseriu-se o *dataset* original (BDo) para uma possível clusterização e separação das classes sem supervisão e, em segundo momento, foi solicitada uma nova clusterização, porém com o banco de dados sintético (BDs). Em ambas as situações, as características apresentadas ao algoritmo foram L*(D65) e PPG-DL, anteriormente selecionadas com o FDR. Os clusters formados com os dados originais são apresentados na Figura 2. Entende-se que o habitual é inserir um banco de dados não catalogado, porém, era desejado avaliar a capacidade de generalização e separação do algoritmo, diante disso, foi avaliado quais são os elementos presentes dentro de cada classe. O mesmo foi realizado para o BDs e os resultados são demonstrados na Figura 2. O *k-means* utilizando um banco de dados menor (BDo) não foi capaz de fazer uma adequada separação das classes, a dificuldade pode estar relacionada ao baixo número de elementos (animais) no banco de dados, há clusters com mais de uma classe presente, embora algumas dessas classes possuam similaridades, esperava-se uma separação mais clara. A composição dos agrupamentos está descrita na Tabela 1.

Apesar dos agrupamentos do BDs também possuírem, em alguns casos, elementos de mais de uma classe, eles são compostos majoritariamente de uma única classe, deixando evidente que para uma adequada separação dos grupos, basta um bom levantamento amostral, tanto em número de elementos quanto da seleção adequada das características descritoras. Ressalta-se que a vantagem de utilizar *k-means* no projeto é seu baixo custo computacional, pois na fase operacional o método se limita ao cálculo de apenas 5 distâncias euclidianas do evento a ser classificado até os 5 centroides definidos pelo *k-means*.

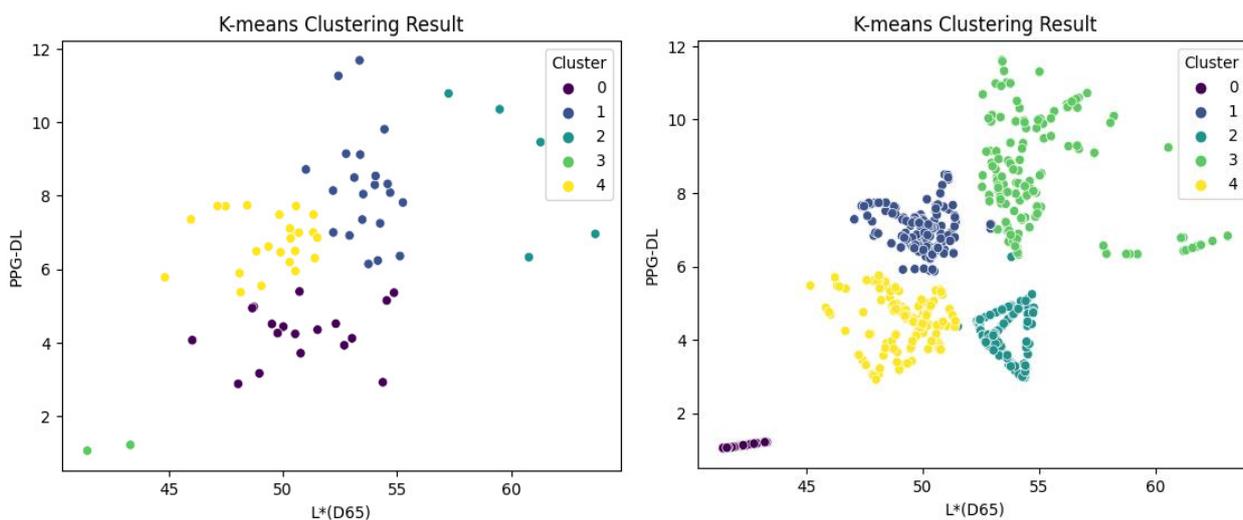


Figura 2. Clusters formados com BDo (a esquerda) e com BDs (a direita).

Tabela 1. Agrupamento *K-means* com BDo e BDs

Clusters	Classes	DFD	RFN	RSE	PFN	PSE
Dados originais (BDo)						
Cluster 0		0	12	0	6	0
Cluster 1		0	0	1	0	20
Cluster 2		0	0	0	0	5
Cluster 3		2	0	0	0	0
Cluster 4		0	5	19	0	0
Dados sintéticos (BDs)						
Cluster 0		150	0	0	0	0
Cluster 1		0	6	150	0	4
Cluster 2		0	1	0	150	1
Cluster 3		0	0	0	0	145
Cluster 4		0	143	0	0	0

Legenda: frequência absoluta de animais de cada classe em cada cluster

CONCLUSÃO

O presente estudo utilizou abordagens de reconhecimento de padrões na análise de qualidade de carnes suínas, empregando Função de Discriminação de Fisher (FDR) e *K-means*. Os resultados obtidos no FDR demonstraram que as características L^* e *drip loss* (PPG-DL) desempenham um papel crucial na discriminação das cinco classes de qualidade suínas (DFD, RFN, SER, PFN e PSE), não apenas corroborando a importância previamente documentada na literatura dessas características, como também quantificou a significância delas em cada classe. O algoritmo *K-means* também foi avaliado, mas não foi capaz de separar as classes adequadamente com o banco de dados original. No entanto, com o aumento do número de amostras, o algoritmo foi capaz de formar clusters com uma única classe dominante. Os resultados deste trabalho mostram o potencial das técnicas de reconhecimento de padrões para a classificação da qualidade da carne suína. Sugere-se para trabalhos futuros, o uso da técnica KNN, que dispensa o treinamento e é considerado uma técnica “zero-shot-learning”. Essas técnicas podem ser usadas para melhorar a eficiência e a qualidade dos processos de produção e comercialização da carne.

AGRADECIMENTOS

À UFPA e as agências de fomento CNPq/Fapemig/CAPES pelo apoio financeiro à pesquisa.

REFERÊNCIAS

- BARBIN, Douglas *et al.* Near-infrared hyperspectral imaging for grading and classification of pork. *Meat Science*, [S.L.], v. 90, n. 1, p. 259-268, jan. 2012. Elsevier BV. <http://dx.doi.org/10.1016/j.meatsci.2011.07.011>.
- CHAWLA, N. V. *et al.* SMOTE: synthetic minority over-sampling technique. *Journal Of Artificial Intelligence Research*, [S.L.], v. 16, p. 321-357, 1 jun. 2002. AI Access Foundation. <http://dx.doi.org/10.1613/jair.953>.
- CHEN, Jung-Chieh; LI, Chih-Peng. Tone Reservation Using Near-Optimal Peak Reduction Tone Set Selection Algorithm for PAPR Reduction in OFDM Systems. *Ieee Signal Processing Letters*, [S.L.], v. 17, n. 11, p. 933-936, nov. 2010. Institute of Electrical and Electronics Engineers (IEEE). <http://dx.doi.org/10.1109/lsp.2010.2077278>.
- FAUCITANO, L. *et al.* Shelf life of pork from five different quality classes. *Meat Science*, [S.L.], v. 84, n. 3, p. 466-469, mar. 2010. Elsevier BV. <http://dx.doi.org/10.1016/j.meatsci.2009.09.017>.
- JAIN, Anil K. *et al.* Data clustering: 50 years beyond *k-means*. *Pattern Recognition Letters*, [S.L.], v. 31, n. 8, p. 651-666, jun. 2010. Elsevier BV. <http://dx.doi.org/10.1016/j.patrec.2009.09.011>.
- JOO, S. T. *et al.* Objectively Predicting Ultimate Quality of Post-Rigor Pork Musculature: ii. practical classification method on the cutting-line. *Asian-Australasian Journal Of Animal Sciences*, [S.L.], v. 13, n. 1, p. 77-85, 1 jan. 2000. Asian Australasian Association of Animal Production Societies. <http://dx.doi.org/10.5713/ajas.2000.77>.
- KAUFFMAN, R.G. *et al.* The effectiveness of examining early post-mortem musculature to predict ultimate pork quality. *Meat Science*, [S.L.], v. 34, n. 3, p. 283-300, jan. 1993. Elsevier BV. [http://dx.doi.org/10.1016/0309-1740\(93\)90078-v](http://dx.doi.org/10.1016/0309-1740(93)90078-v).
- MACQUEEN, J. (1967). Some methods for classification and analysis of multivariate observations. In *Proceedings of the fifth Berkeley symposium on mathematical statistics and probability*, 1(14), 281-297.
- PERES, Louise Manha *et al.* Fuzzy approach for classification of pork into quality grades: coping with unclassifiable samples. *Computers And Electronics In Agriculture*, [S.L.], v. 150, p. 455-464, jul. 2018. Elsevier BV. <http://dx.doi.org/10.1016/j.compag.2018.05.009>.
- SILVA, Rebeca N. *et al.* Non-invasive method to analyse the risk of developing diabetic foot. *Healthcare Technology Letters*, [S.L.], v. 1, n. 4, p. 109-113, out. 2014. Institution of Engineering and Technology (IET). <http://dx.doi.org/10.1049/htl.2014.0076>.
- THEODORIDIS, Sergios; KOUTROUMBAS, Konstantinos. *Pattern recognition*. Elsevier, 2006.
- TOMOVIC, Vladimir Milo *et al.* Technological quality and composition of the M. semimembranosus and M. longissimus dorsi from Large White and Landrace Pigs. *Agricultural And Food Science*, [S.L.], v. 23, n. 1, p. 9-18, 25 fev. 2014. Agricultural and Food Science. <http://dx.doi.org/10.23986/afsci.8577>.
- TORRES FILHO, Robledo de Almeida *et al.* Classification of pork quality by hierarchical cluster analysis. *British Food Journal*, [S.L.], v. 120, n. 7, p. 1446-1456, 2 jul. 2018. Emerald. <http://dx.doi.org/10.1108/bfj-10-2017-0589>.
- VAN LAACK, R L *et al.* Glycolytic potential of red, soft, exudative pork Longissimus muscle. *Journal Of Animal Science*, [S.L.], v. 77, n. 11, p. 2971, 1999. Oxford University Press (OUP). <http://dx.doi.org/10.2527/1999.77112971x>.
- WARNER, R.D. *et al.* Muscle protein changes *post mortem* in relation to pork quality traits. *Meat Science*, [S.L.], v. 45, n. 3, p. 339-352, mar. 1997. Elsevier BV. [http://dx.doi.org/10.1016/s0309-1740\(96\)00116-7](http://dx.doi.org/10.1016/s0309-1740(96)00116-7).
- XU, Dongkuan *et al.* A Comprehensive Survey of Clustering Algorithms. *Annals Of Data Science*, [S.L.], v. 2, n. 2, p. 165-193, jun. 2015. Springer Science and Business Media LLC. <http://dx.doi.org/10.1007/s40745-015-0040-1>.