MEAT TENDERNESS PREDICTION USING MACHINE LEARNING: DISTRIBUTION APPLIED TO VARIABLES

Gabriel Furini[1], Angelo Polizel Neto[1], Heinsten Frederich Leal dos Santos [1]

[1]*Federal University of Rondonopolis*
5055, Student Avenue, Sagrada Familia, CEP 78736900 - Rondonopolis - Mato Grosso State, Brazil
gabriel.furini@aluno.ufr.edu.br, polizel.neto@ufr.edu.br, heinsten.leal@ufr.edu.br

Abstract. Beef tenderness is an attribute of great prestige for consumers. Several factors, such as genetics, food, and environment, influence meat tenderness, which can be objectively evaluated after the animal is slaughtered. Typically, sensitivity measurement involves mechanical testing, in which the shear force required to break muscle fibers is quantified. This study aims to validate a more comprehensive and robust database, incorporating hyperspectral imaging, to help predict meat tenderness non-destructively. In this new approach, computational techniques, such as machine learning with the use of artificial neural networks, are employed to explore the dependence of transmitted variables without the need to control the response. Hyperspectral images provide information about specific wavelengths, allowing for a more detailed analysis of the sample. To assess tenderness parameters, measures such as pH, sample color, hot carcass weight, ribeye area, breed, and sex, as well as hyperspectral images, along with fillet shear force. The objective is to identify the relationship between these variables and the shear force required to break the muscle fibers. The use of normal and gamma mathematical distributions as tools for more comprehensive training in Machine Learning was used. The analysis showed that the statistical model adequately adjusted the data, confirming the reliability and accuracy of the forecasts obtained. This additional validation reinforces the robustness of the proposed model, which can handle the variability of beef tenderness data. With these new arrangements, it was possible to study the behavior of these distributions through validation from a robust database in which Random Forest algorithms were implemented in Machine Learning and Neural Network. Based on the data presented, a Coefficient of Determination (R2) of 0.4494 was obtained, demonstrating the effectiveness of the Machine Learning model in predicting the shear force in relation to the values obtained in the mechanical tests. This innovative approach, using hyperspectral imaging in conjunction with machine learning techniques, provides a broader and more robust database for predicting beef tenderness. These scientific advances have the potential to improve end-product quality and meet consumer expectations for beef tenderness.

Keywords: Machine Learning, Computational Methods, Prediction, Tenderness, Distribution.

1 Introduction

Brazil's significant growth trajectory has positioned it as a global leader in beef production, boasting the world's largest commercial bovine herd at a staggering 214.7 million cattle. This thriving industry sees an annual throughput of 44 million carcasses, as reported by the National Confederation of Agriculture ([1] and [2]). Notably, the state of Mato Grosso stands as the chief contributor to this meat production juggernaut, housing a herd exceeding 30.9 million cattle. Impressively, its meat exports surged by 15.37% in 2020, amounting to 47.95 thousand tonnes equivalent of carcass sold ([3]). With a responsibility to nourish a population of 210 million, Brazil's beef sector caters to numerous specialized markets, capturing a substantial 80% of the domestic market. This translates into a revenue of R$97.3 billion, while exports to over 150 countries rake in a further R$24.1 billion. This global market share accounts for 22% of the world's beef exports ([4], [1], [2]).

In this context, the Brazilian beef industry's aspiration to be a premier exporter hinges on the quality of its products. Quality encompasses an extensive spectrum of attributes, spanning intrinsic and extrinsic qualities. Among these, the paramount significance of meat tenderness emerges, as unequivocally stated by [5]. This attribute profoundly influences consumer acceptance, satisfaction, and repurchase rates. Key determinants of meat tenderness include the amount and solubility of connective tissue, the composition and state of muscle fibers, and the degree of muscle proteolysis ([6]). The Shear Force measurement, introduced by Bratzler in 1954 and standardized by the American Meat Science Association in 1995, stands as the gold standard for evaluating meat tenderness ([7]).

Anticipating and optimizing meat tenderness involves assessing variables such as color, age, sex, and species, which collectively dictate visual appeal and sensory acceptability. These attributes hold economic implications, as undesirable color and weight loss lead to financial losses. Additionally, the industry can command premium prices by guaranteeing tenderness and eating quality ([7]). The intricate interplay of biochemical properties and pH levels in meat is governed by pre- and postmortem influences, which in turn impact muscle cell structure and connective tissue.

The essence of pH level shifts lies in their effect on protein denaturation, the spacing of myofibrillar lattices, and muscle cell contraction, collectively shaping both raw and cooked meat quality. The field of meat science converges on two pivotal research avenues, driven by the pursuit of safe, high-quality meat catering to a diverse consumer base. Among the triad of flavor, juiciness, and, foremost, tenderness, it is the latter that most dominantly shapes consumer perceptions ([8]). Advancements over the years have pinpointed the key variables underpinning tender meat ([9]).

The quest for objective measures of beef quality has spurred extensive research, culminating in the development of tools for industry application. Notably, the Warner–Bratzler shear force instrument has gained prominence, as its measured shear strength correlates with sensory tenderness scores ([10]). Recent years have witnessed heightened interest in swiftly implementable tools for the food industry, fostering the exploration of modern techniques such as electronic noses, computer vision, spectroscopy, and spectral imaging. These methods extract copious digital data on food attributes, necessitating robust data analysis due to redundancy and irrelevance. To tackle this, advanced techniques like artificial neural networks (ANN), random forest, k-nearest neighbor (KNN), and other computational approaches have emerged as valuable tools ([11]).

Artificial intelligence techniques, particularly neural networks (ANN), have emerged as a cornerstone for evaluating subjective attributes like tenderness ([12]). Studies have demonstrated the application of ANN for meat quality assessment, utilizing simple physical measurements to fuel a supervised learning approach. This drive towards prediction proficiency signals a growing interest in leveraging artificial intelligence for meat quality and composition prediction, ultimately shaping the trajectory of the industry ([12]).

This paper apply into the methodology of mathematical distributions and their consequential impact on predictive recognition using machine learning and neural networks. The focal objective here is to underscore the pivotal significance of specific distribution types in yielding enhanced values for prediction or recognition patterns within the context of machine learning.

The investigation is further characterized by its keen emphasis on the discernment of feature importances inherent within these distributions, thereby furnishing invaluable insights conducive to the optimization of the predictive models.

The interplay between distribution profiles and their ensuing outcomes, the current study significantly augments our comprehension of the pivotal role played by specific mathematical properties in elevating the accuracy of prediction and recognition frameworks. This contribution advances the discourse surrounding the nuanced interplay of distribution types in optimizing the efficacy of machine learning applications, particularly in the realm of predictive modeling.

2 Methodology

The foundational framework for evaluating meat tenderness is established upon an expansive dataset derived from a shear force methodology applied to meat samples. The study now encompasses a dataset of 612 meat samples, meticulously divided into uniform dimensions of 10x10x30 mm, aligned along the fiber orientation. Employing the mean maximum force ((given as $N.cm^{-2}$)), the subsequent data analysis encompasses a suite of attributes for each sample, spanning age, animal gender, pH levels, shear force, and data from hyperspectral images. This dataset, cultivated within a controlled laboratory milieu, subsequently undergoes computational analysis through two principal methodologies: Random Forest Machine Learning and Neural Network, with the overarching objective of predicting the shear force values ascertained within the laboratory setting.

The Random Forest approach, a notable machine learning paradigm pioneered by Breiman in 2001 [13], entails an assembly of classification and regression trees [13]. These arboreal structures, founded on binary variable subdivisions to deduce predictive outcomes, collectively amalgamate their outputs through a majority consensus, culminating in a final decision. The trees are constructed through bootstrapped replicas of the original dataset, with a set of entry features being randomly selected from the available feature pool (F) at each tree node. Among these features, the most optimal division is identified, underscoring the notable efficacy of Random Forests that supersedes many alternative classifiers [14], while concurrently exhibiting robustness against overfitting.

Artificial Neural Networks (ANN), a diverse range of computational techniques emulating biological neural structures, engender mathematical models proficient in optimizing performance through iterative training and applications (Araujo et al. [15]). These networks are characterized by interconnected processing nodes, epitomizing the capacity to amass knowledge and generalize from training data. Mirroring human cognitive capabilities, neural networks excel in deciphering intricate patterns, in contradistinction to conventional computing paradigms, which excel in algorithmic computational tasks.

Within the purview of this study, Machine Learning is grounded in a dataset of 612 samples, each characterized by 259 variables sourced from the laboratory milieu. Employing the Random Forest Regressor approach, the dataset is partitioned into training and validation subsets at a ratio of 60% to 40%, culminating in an accuracy of 85%. Importantly, the relative significance of each variable within the network training process is quantified, offering discernment into their respective contributions. Subsequent evaluation extends to a dataset of 350,000 samples, (https://drive.google.com).characterized by four variables (PH, Animal Sex, Age, and Shear Force), in conjunction with an array of spectral bands procured through spectrophotometric measurements. The model is subjected to rigorous scrutiny, yielding accuracy metrics spanning from 73.72% to 75.06%, affirming its adeptness in discerning patterns within the dynamic Gamma and Normal distributions that encapsulate the evolving analytical approach.

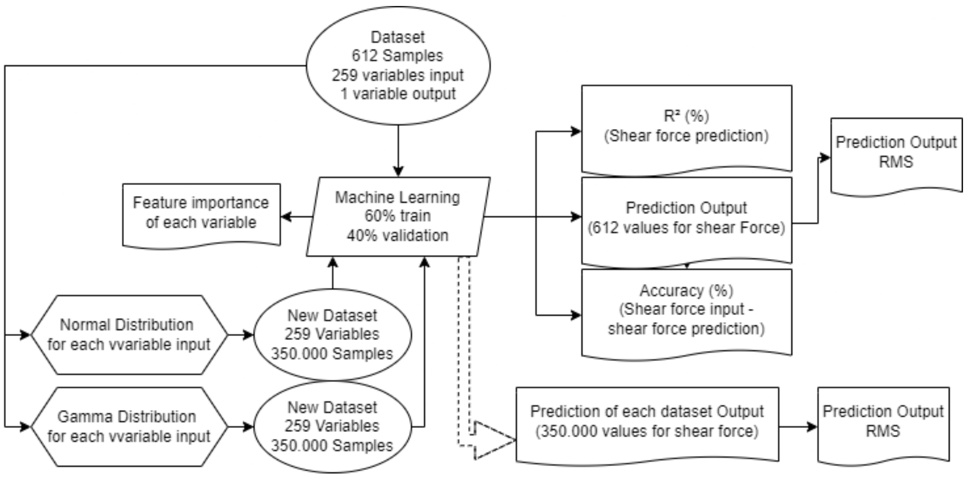

Figure 1. Evaluation Workflow for Machine Learning Prediction

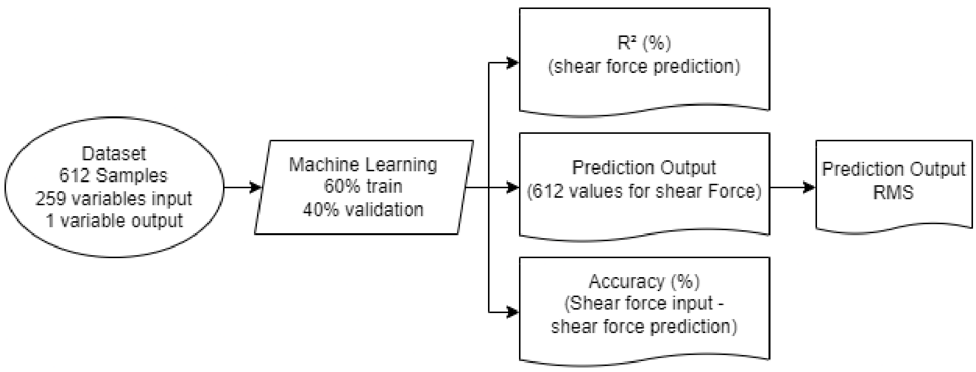

Figure 2. Evaluation Workflow for Neural Network

2.1 Results

In this section, we present the outcomes obtained through the utilization of Machine Learning and Neural Network methodologies, alongside a mathematical procedure employed to establish novel databases for the variables under consideration. To create two fresh databases per variable, we employed both a normal distribution and a gamma distribution. These distributions were configured to share identical mean and standard deviation values, ensuring the robustness of the data.

Furthermore, Machine Learning illustrated the importance of each variable in the prediction process through

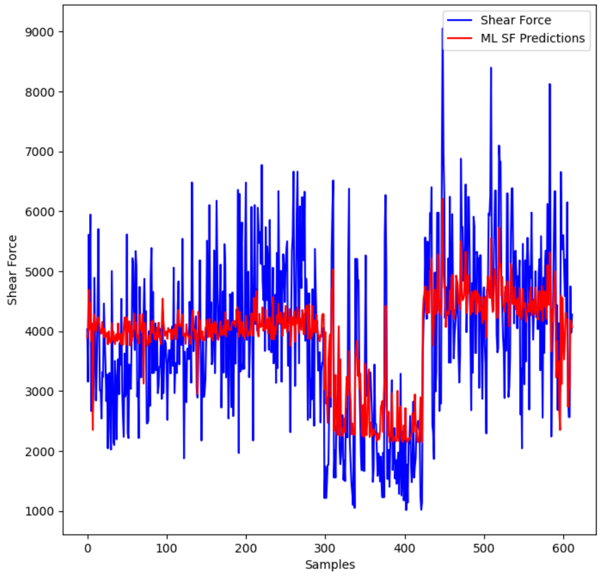

Figure 3. Comparison Values of Shear Force Measure vs Prediction

a tree-like visual representation. The analysis highlighted the following variable significance percentages: pH: 31.10%, VAR3: 6.30%, VAR2: 5.94%, Animal_Sex: 4.96%, and VAR5: 2.82%. To prognosticate shear force values for the newly introduced inputs, we leveraged a well-honed machine learning framework. The RMS was subsequently calculated for the fresh Shear Force data, exhibiting conformance to the predefined parameters and displayed values. This congruity in outcomes underscores the mutual efficacy of both the Machine Learning and Neural Network approaches, thereby inviting a thorough and insightful comparison.

3 Conclusions

Contemporary progressions within computational methods and the domain of meat sciences have engendered avenues for anticipating meat quality and identifying pivotal variables in gauging meat tenderness. Within this context, this study showcases the efficacy of computational approaches in prognosticating shear force values, achieving an accuracy rate of 75.06%. This achievement is rooted in the predicted machine-learning values, derived from an extensive database of 612 samples. By capitalizing on genuine shear force values extracted from the array of 612 meat samples, the ensuing Root Mean Square (RMS) value of 63.91 emerges as a foundational benchmark. This RMS figure offers a pivotal frame of reference, facilitating a comprehensive evaluation and contrast across various prediction methodologies.((Table 1)

Table 1. Comparison of Root Mean Square (RMS) Values for Shear Force Prediction

Methodology	**RMS Value**
Genuine Shear Force (612 samples)	63.91
Machine Learning (Original 612 samples)	62.41
Machine Learning (350,000 samples - Normal Distribution)	64.22
Machine Learning (350,000 samples - Gamma Distribution)	56.70
Neural Network (Post-training with 612 samples)	59.61

This comprehensive study emphasizes the predictive potential of computational techniques in discerning shear force values. The outcomes effectively confirm the efficacy of both Machine Learning and Neural Network approaches, along with the extended subsets. Additionally, the coefficient of determination ($R^2 = 49.66\%$) serves as evidence of the consistent alignment between predicted and actual shear force values, with Root Mean Square (RMS) employed as a complementary reference.

In anticipation of future undertakings, this research trajectory will entail an enriched analysis involving an augmented dataset, housing authentic shear force measurements. This expansion is poised to refine the predictive models further, with the primary aim of enhancing the coefficient of determination (R^2) by leveraging an augmented Real dataset parameter. The discernible relationship established between the generated dataset and precision portends significant potential. The integration of genuine measurements is anticipated to precipitate a reconfiguration of the Machine Learning framework, offering enhanced adaptability and accuracy. With real-world data assimilated, the models are poised for heightened predictive efficacy, ultimately augmenting their relevance within the realm of meat quality assessment. This iterative refinement process holds promise for elevating predictive models towards heightened precision and robustness, thereby fostering more accurate and insightful evaluations within the expansive domain of meat sciences.

Acknowledgements. We would like to extend our sincere gratitude to FAPEMAT (Foundation for Research Support of the State of Mato Grosso) for their invaluable support in facilitating the development of this paper. The collaborative efforts of the Research Group SmartLabs have been instrumental in shaping the trajectory of this research, and we acknowledge their dedication and contributions with utmost appreciation. Additionally, we are indebted to the Federal University of Rondonopolis for providing the necessary resources and conducive environment that enabled the fruition of this study. This work would not have been possible without the combined efforts and support of these institutions, which have played a vital role in the advancement of our research endeavors.

Authorship statement. The authors here by confirm that they are the sole liable persons responsible for the authorship of this work, and that all material that has been herein included as part of the present paper is either the property (and authorship) of the authors, or has the permission of the owners to be included here.

References

[1] CNA. *Confederação da Agricultura e Pecuária do Brasil*. cnabrasil, https://www.cnabrasil.org.br/noticias/cna-apresenta-balanco-de-2020-e-as-perspectivas-para-2021, 2021.

[2] CNA. *Confederação da Agricultura e Pecuária do Brasil*. cnabrasil, https://www.cnabrasil.org.br/cna/panorama-do-agro, 2020.

[3] IMAC. *Exportações de carne bovina de MT crescem 15% em outubro*. IMAC, https://www.imac.agr.br, 2020.

[4] da Cadeia Produtiva da Carne Bovina Brasileira: Uma Visão para O. F. 2040. *Exportações de carne bovina de MT crescem 15% em outubro*. EMBRAPA, https://www.embrapa.br, 2020.

[5] e. a. M.F.. Miller. Consumer thresholds for establishing the value of beef tenderness. *Journal of animal Science*, vol. 79, pp. 3062 – 3068, 2001.

[6] S. e. a. Joo. Control of fresh meat quality through manipulation of muscle fiber characteristics. *Meat Science*, vol. , 2013.

[7] e. a. Monteiro. identification of the marble meat standard through the neural network for quality grading. *25th ABCM International Congress of Mechanical Engineering*, vol. , 2019.

[8] F. L. Z. Q. Chu Zhang. Lei Zhou and Y. He. Application of deep learning in food: A review. *Comprehensive Reviews in Food Science and Food Safety*, vol. 0, 2019.

[9] P. P. R. S.K. Oiseth b. J.M. Hughes a. A structural approach to understanding the interactions between colour, water-holding capacity and tenderness. *Meat Science*, vol. , 2014.

[10] I. M. F. Member. Juan M. Fontana and I. Senior Member. Edward Sazonov*. Estimation of feature importance for food intake detection based on random forests classification. *35th Annual International Conference of the IEEE EMBS*, vol. , pp. 2–4, 2013.

[11] C. C. Ali Seydi Keçeli. Aydin Kaya and B. Tekinerdogan. Sensor failure tolerable machine learning-based food quality prediction model. *Sensors*, vol. , 2020.

[12] C. Lin and C. Lee. Neural-network-based fuzzy logic control and decision system. *IEEE Trans. Comput*, vol. 40, pp. 1320–1336, 1991.

[13] A. Géron. *Hands-on Machine Learning with Scikit-Learn, Keras TensorFlow*. O'Reilly Media, Inc., 1005 Gravenstein Highway North, Sebastopol, CA 95472., 2019.

[14] e. a. Devin A. Gredell1. Comparison of machine learning algorithms for predictive modeling of beef attributes using rapid evaporative ionization mass spectrometry (reims) data. *Scientific Reports*, vol. 0, n. 0, pp. 1–7, 2019.

[15] V. S. Araujo, G. S. Silva, and H. F. L. Santos. Shunt control on smart strucutres using ge/netic algorithm and neural network method. Buzios, Brazil, 2019.