International Journal of Agricultural Technology 2024 Vol. 20(6):2497-2512
Available online http://www.ijat-aatsea.com
ISSN 2630-0192 (Online)

Using orange data mining for meat classification: The preliminary application of machine learning

Phoemchalard, C.[1,2] [*], Senarath, N.[1,2], Malila, P.[1,2], Tathong, T.[3] and Khamhan, S.[4]

[1]Department of Agriculture, Mahidol University, Amnatcharoen Campus, Amnatcharoen 37000, Thailand; [2]Excellence Center on Agriculture and Food for Health, Mahidol University, Amnatcharoen Campus, Amnatcharoen 37000, Thailand; [3]Department of Food Technology, Faculty of Agriculture and Technology, Nakhon Phanom University, Nakhon Phanom 48000, Thailand; [4]That Phanom College, Nakhon Phanom University, Nakhon Phanom 48110, Thailand

Phoemchalard, C., Senarath, N., Malila, P., Tathong, T. and Khamhan, S. (2024). Using orange data mining for meat classification: The preliminary application of machine learning. International Journal of Agricultural Technology 20(6):2497-2512.

Abstract Orange Data Mining study on the classification of buffalo, beef, and goat meats, Machine Learning (ML) classifiers including Support Vector Machine (SVM), Neural Network (NN), and Naïve Bayes (NB) are well performed to achieve 100% accuracy across all features. Random Forest (RF) demonstrated the best performance more than 97% in AUC, CA, F1, and MCC. Other models such as Gradient Boosting (GB), AdaBoost, CN2 Rule Induction (CN2), Decision Tree (DT), and k-Nearest Neighbors (KNN) are performed better but there were less efficient. In the application of specific classifiers for species-based meat quality attributes, SVM, NN and NB should be considered as the best options.

Keywords: Machine learning, Data mining, Classification algorithms, Meat quality, Ruminants

Introduction

Machine learning (ML) may be considered a crucial meeting point in computer science, artificial intelligence, mathematics, and statistics (Arora, 2018; Dayal *et al*., 2023; Duarte and Ståhl, 2019). ML focuses on the creation of advanced algorithms and models that enable computer systems to learn and improve data autonomously without the need for explicit programming in every situation (Kumar and Hasija, 2021). ML comprises methods such as mathematical optimization, statistical analysis, and algorithm training on big datasets, enabling machines to obtain similar knowledge and capabilities to humans via learning and experience (Head and Aloqaily, 2022; Sharma and Bharti, 2021; Thakur, 2023).

[*] **Corresponding Author:** Phoemchalard, C.; **Email:** chirasak.pho@mahidol.edu

The two main approaches to ML are supervised and unsupervised learning; entirely opposite in terms of thinking. Supervised learning uses labeled data and human supervision to train models for jobs such as prediction and classification (Gupta *et al*., 2022). The most commonly used supervised algorithms are linear regression, logistic regression, Naïve Bayes, k-Nearest Neighbor, Support Vector Machines, Random Forest, and Convolutional Neural Networks (Rashidi *et al*., 2019). Supervised learning seeks to find rules and relationships using labeled data. In contrast, unsupervised learning provides the basis for learning using unlabeled data to derive the inherent patterns, trends, and structure of the data itself (Elsaadouny *et al*., 2021; Sen and Das, 2023). Although these two concepts offer robust implementation of feature models, the first estimation approach is superior in terms of forecasting fields.

The range of ML applications is broad and growing constantly (Jordan and Mitchell, 2015; Peng *et al*., 2021), for example, improving medical image analysis for medical diagnosis or serving natural language processing systems to increase customer experiences. As the world generates increasing volumes of complex data, ML drives innovations in processing and extracting valuable insights across domains (Sathya *et al*., 2022; Sharma and Bharti, 2021). ML will be the leading force in the development of AI up to the point at which AI can perform human cognitive functions. It runs intelligent systems for improving processes, identifying patterns, and making data-based decisions with real consequences (Head and Aloqaily, 2022; Jordan and Mitchell, 2015). This multidisciplinary research is at the forefront of computer science and the incorporation of technology innovation (Domingos, 2012; Jordan and Mitchell, 2015).

The k-Nearest Neighbors algorithm transforms meat classification using technologies such as electrical impedance spectroscopy, multispectral imaging, and computer vision, which are fed into algorithms like Support Vector Machines and Convolutional Neural Networks. These AIs precisely distinguish meat types, recognize adulterants, and predict quality traits like freshness, tenderness, and color through the non-destructive, non-invasive analysis of images, spectra, and sensor signals (Ayaz *et al*., 2020; Li *et al*., 2021a; Malikhah *et al*., 2021; Medeiros *et al*., 2021; Qiu *et al*., 2024; Setiadi *et al*., 2022; Torres *et al*., 2022; Yang *et al*., 2023). ML is a powerful tool for performing automated, unbiased inspections. Reliable meat classification, grading, and quality control have become a reality because of these improvements. Finally, it is the basis on which food safety and food authenticity in the meat industry depend.

Therefore, the aim of this study was to explore the application of the Orange Data Mining Tool in classifying buffalo, beef, and goat meat using nine machine learning algorithms.

Materials and methods

Sample collection and physicochemical analysis

Samples from *longissimus et lumborum* muscle were collected from 40 buffaloes, 30 beef cattle, and 32 goats at local abattoirs in Amnat Charoen and Nakhon Phanom Provinces. Two kilograms of carabeef and beef and one kilogram of goat meat sample were purchased, packed in zip-lock bags, stored at 4°C, and transported immediately to the meat laboratory. After unpacking, the samples were cut, trimmed, and subjected to quality parameter analyses.

Meat quality parameters including pH, color, water-holding capacity, texture, chemical composition and odor were analyzed. The pH of the meat was determined in triplicate using a pH meter equipped with a stainless-steel probe (HI99163, Hanna Instruments, Inc., Woonsocket, RI, USA). Instrumental color measurements were performed in quintuplicate on the cut surface using a colorimeter (CR-400, Konica Minolta, Tokyo, Japan), with CIE L*a*b* values recorded to measure lightness (L*), redness (a*), and yellowness (b*) according to standard AMSA recommendations (AMSA, 2012). Drip and cooking loss measurements were used to assess water-holding capacity (Honikel, 1998). A texture analyzer (Stable Micro Systems Ltd., Surrey, UK) was used to measure the Warner-Bratzler shear force (AMSA, 2016) and the texture profile (Bourne, 1978) of cooked meat as described by (Phoemchalard *et al*., 2022). Furthermore, the proximate composition, such as moisture, protein, fat, and ash, was determined using established protocols (AOAC, 2012). Finally, an electronic nose system with eight metal oxide sensors was used to assess the volatile flavor profiles of a set of meat samples (Phoemchalard *et al*., 2021).

ML algorithms and computer specifications

Nine well-known classifiers were used in this study: Decision Tree (DT), Random Forest (RF), Support Vector Machine (SVM), Gradient Boosting (GB), Neural Network (NN), CN2 Rule Induction (CN2), k-Nearest Neighbors (KNN), Naïve Bayes (NB), and AdaBoost. Computational analysis was conducted on a personal laptop with a powerful processor (AMD Ryzen 7 4800H), 16 GB RAM, 512 GB SSD, and an NVIDIA GeForce GTX 1650 GPU. Windows 11 was used as the operating system. The Ryzen 7 4800H processor runs at a base clock speed of 2.9 GHz and a peak boost clock of 4.2 GHz. This chip has eight cores and 16 threads, resulting in efficient parallel processing. In accordance with the given indicators, the laptop could process big and complicated datasets quickly. Furthermore, the algorithm could easily generate unique visualizations.

Evaluation metrics

The effectiveness of the classification models in terms of various aspects, such as Area Under the Curve (AUC), Accuracy, Recall, Precision, F1 score, and Matthews Correlation Coefficient (MCC), was estimated using different metrics to determine predictive performance. The principal metrics utilized in the current investigation are delineated as follows:

- The AUROC determines the classifier's precision with a value varying from 0.9–1.0 thought to be exceptional, 0.8–0.9 excellent, 0.7–0.8 reasonable, 0.6–0.7 inadequate, and 0.5–0.6 as failing (Forsyth, 2018)
- Accuracy = (TN + TP) / (TP + TN + FP + FN)
- Recall = TP / (TP + FN)
- Precision = TP / (TP + FP)
- F1 score = 2 * ((Precision * Recall) / (Precision + Recall))
- MCC = (TP * TN - FP * FN) / sqrt ((TP + FP) * (TP + FN) * (TN + FP) * (TN + FN))

Where, TP = True Positive, TN = True Negative, FP = False Positive, and FN = False Negative

Data simulation

The Orange Data Mining tool version 3.36.2 (Demšar *et al*., 2013) was used to conduct a thorough analysis of meat quality parameters, including pH, color characteristics (lightness, redness, and yellowness), water-holding capacity (drip loss and cooking loss), proximate composition (moisture, protein, fat, and ash), textural properties (shear force, hardness, adhesiveness, springiness, cohesiveness, gumminess, and chewiness), and meat odors (eight distinct sensors). The simulation process was organized into specific steps starting with the choice of a 102-meat quality dataset (Figure 1). The relevant features were then identified to predetermine the meat species. As essential parts of the model, data preprocessing techniques were applied. The dataset was then included in a 66-sample training set. Hence, a method of deep learning that relies on historical data was used and predictive models were developed and tested on an unknown dataset. A series of ML algorithms was simulated, and their performance measures were calculated and analyzed.

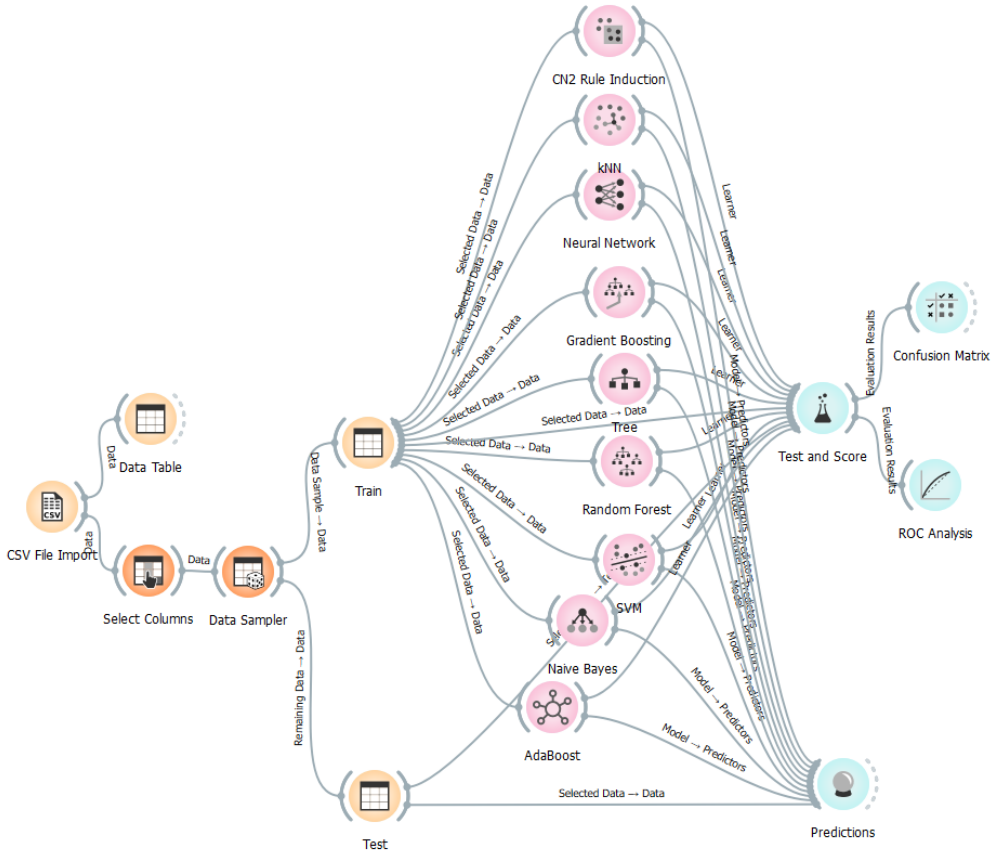

Figure 1. Workflow of the classification model using Orange Data Mining

Results

In this study, the selected algorithms were trained and then tested on the training and test datasets. The values of the key performance indicators, i.e., AUC, CA, F1 score, Precision, Recall, and MCC, for all classification methods used on the training data are shown in Table 1 and Figure 3. With these metrics, data on the accuracy of the models during the training phase were obtained regarding the prediction of target variables. The results revealed that the Support Vector Machine, Neural Network, and Naïve Bayes classifiers were the most perfect in classifying meat types (100%). The Random Forest and Gradient Boosting classes obtained the highest scores for all criteria. Thus, their performance was the most accurate on average. The Random Forest model attained an impressive benchmark AUC value of 0.988, providing the best possible result. Moreover, it achieved an accuracy rate of 0.917 and an F1 value of 0.916, indicating remarkable accuracy and promising results. The Gradient

Boosting classifier had an AUC of 0.965, indicating the high accuracy of this model. Furthermore, CA demonstrated high accuracy (0.889), while the F1 score also showed a robust correlation (0.888). The performance of this model is expected to be very close to that of other models, making it very competitive.

The Tree, CN2 Rule Induction, k-Nearest Neighbors, and Ada Boost Classifiers showed a different performance for each metric. Some models worked better on certain criteria, while others were unsuitable. The Tree classifier obtained a rating of 0.891, indicating good performance. However, the CA and F1 scores were only 0.806 and 0.801, implying that it may not be the best choice. Regarding CN2 Rule Induction, this classifier attained an AUC of 0.912, suggesting excellent efficiency. Its CA and F1 ratings were also high (0.861 and 0.862) explicitly making it a practical choice for this task. For the k-Nearest Neighbors, it achieved an AUC of 0.876, indicating moderate performance. Nevertheless, its CA and F1 scores were only 0.681, demonstrating that the equipment may be irrelevant for this purpose.

Table 1. Average performance of classification algorithms on the training dataset

Model	AUC	CA	F1	Prec	Recall	MCC
Tree	0.891	0.806	0.801	0.801	0.806	0.701
Random Forest	0.988	0.917	0.916	0.916	0.917	0.872
Support Vector Machine	1.000	1.000	1.000	1.000	1.000	1.000
Gradient Boosting	0.965	0.889	0.888	0.887	0.889	0.830
Neural Network	1.000	1.000	1.000	1.000	1.000	1.000
CN2 Rule Induction	0.912	0.861	0.862	0.865	0.861	0.787
k-Nearest Neighbors	0.876	0.681	0.681	0.683	0.681	0.510
Naïve Bayes	1.000	1.000	1.000	1.000	1.000	1.000
AdaBoost	0.909	0.875	0.875	0.877	0.875	0.812

AUC = Area Under the Curve, CA = Classification Accuracy, F1 = F1 score, Prec = Precision, Recall = True Positive Rate, and MCC = Matthews Correlation Coefficient

The confusion matrix is a performance evaluation technique for ML classifiers with two or more output classes. It enables the classification models to be compared in terms of performance. For the Support Vector Machine, Neural network, and Naïve Bayes, perfect classification rates of 100% were realized, later resulting in the complete eradication of any misclassification in meat identification (Figure 2). These models are the best option here due to their precise and reliable patterns for this specific situational assignment.

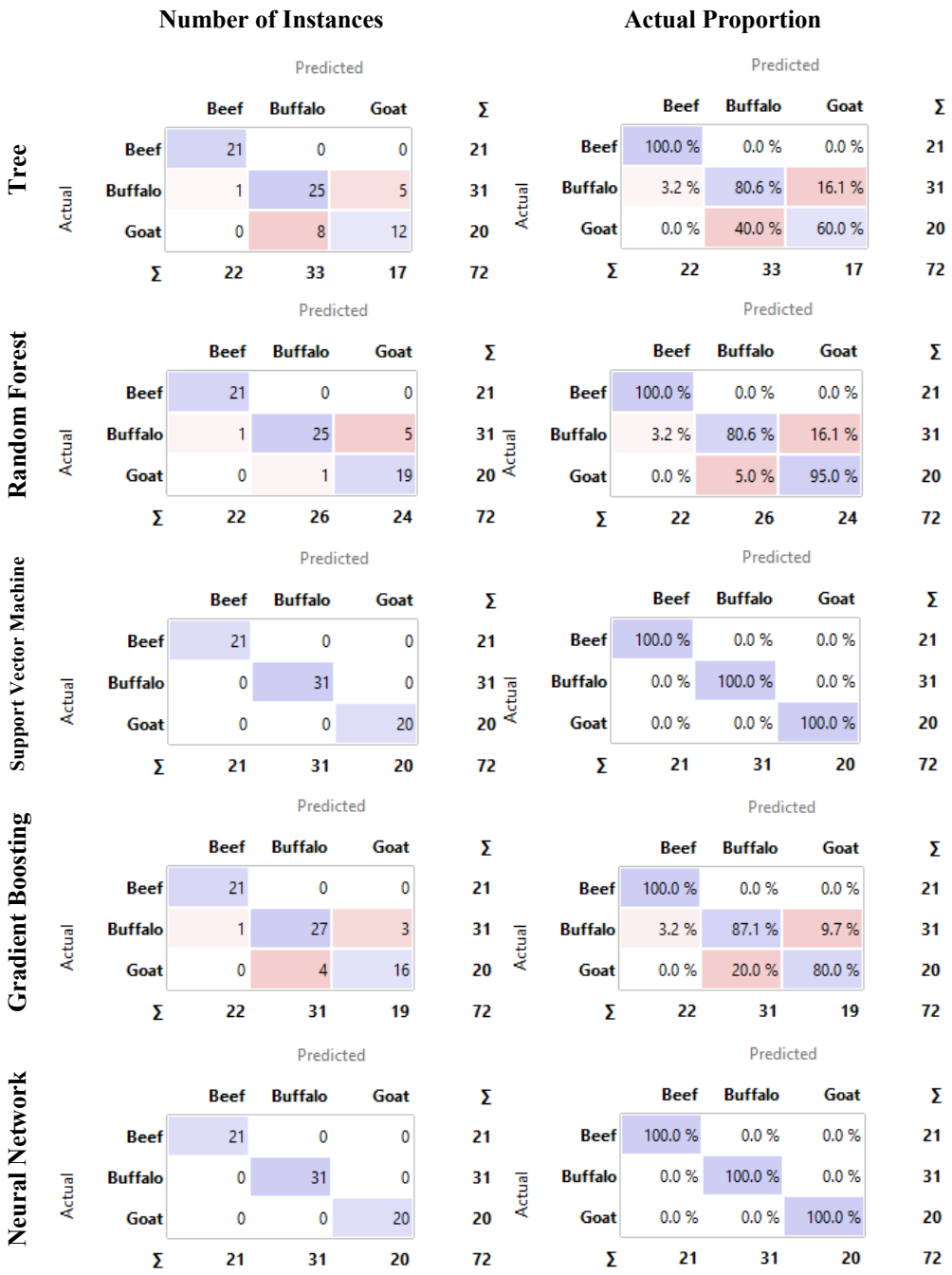

Figure 2. Confusion matrix values of the nine classifier algorithms

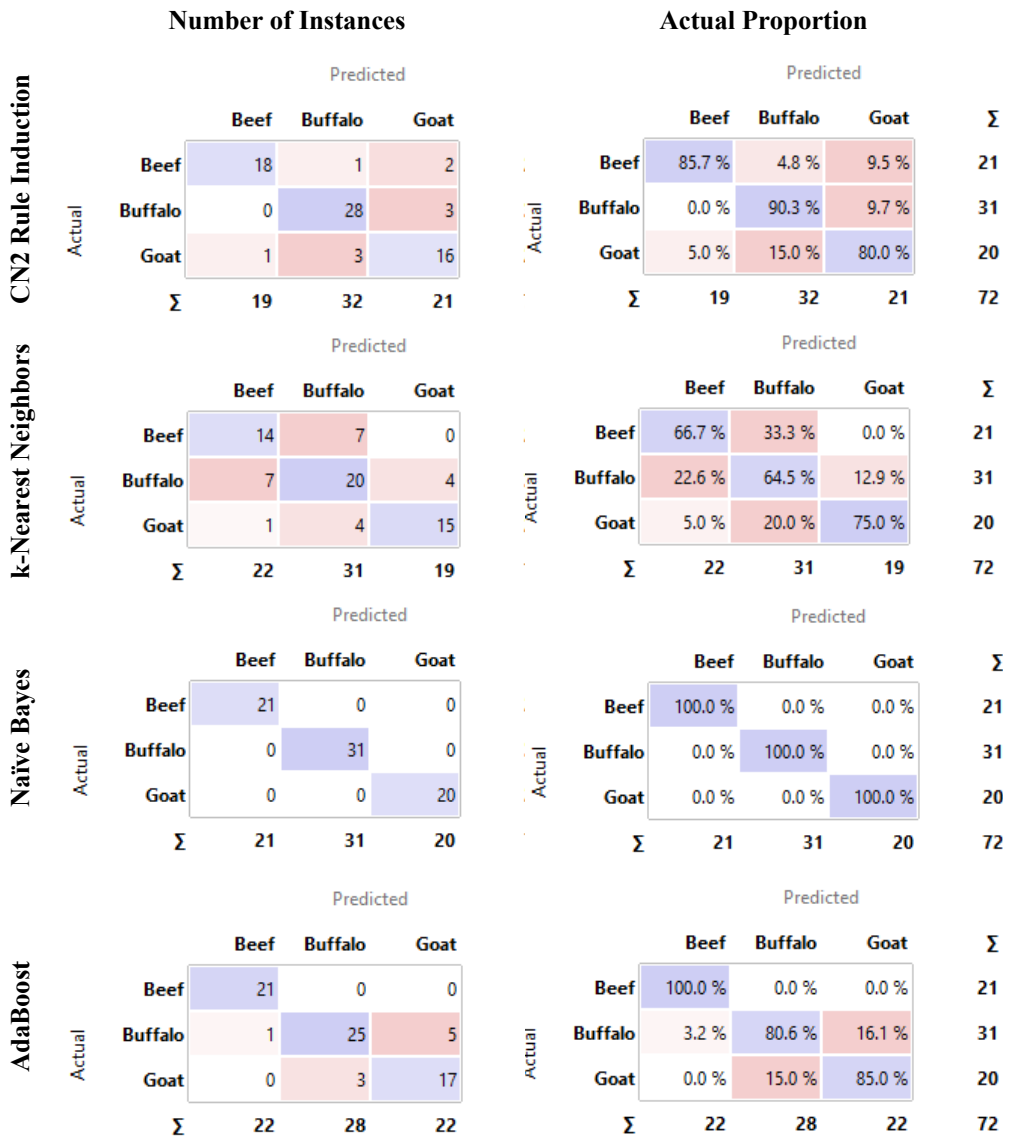

Figure 2. Confusion matrix values of the nine classifier algorithms (cont.)

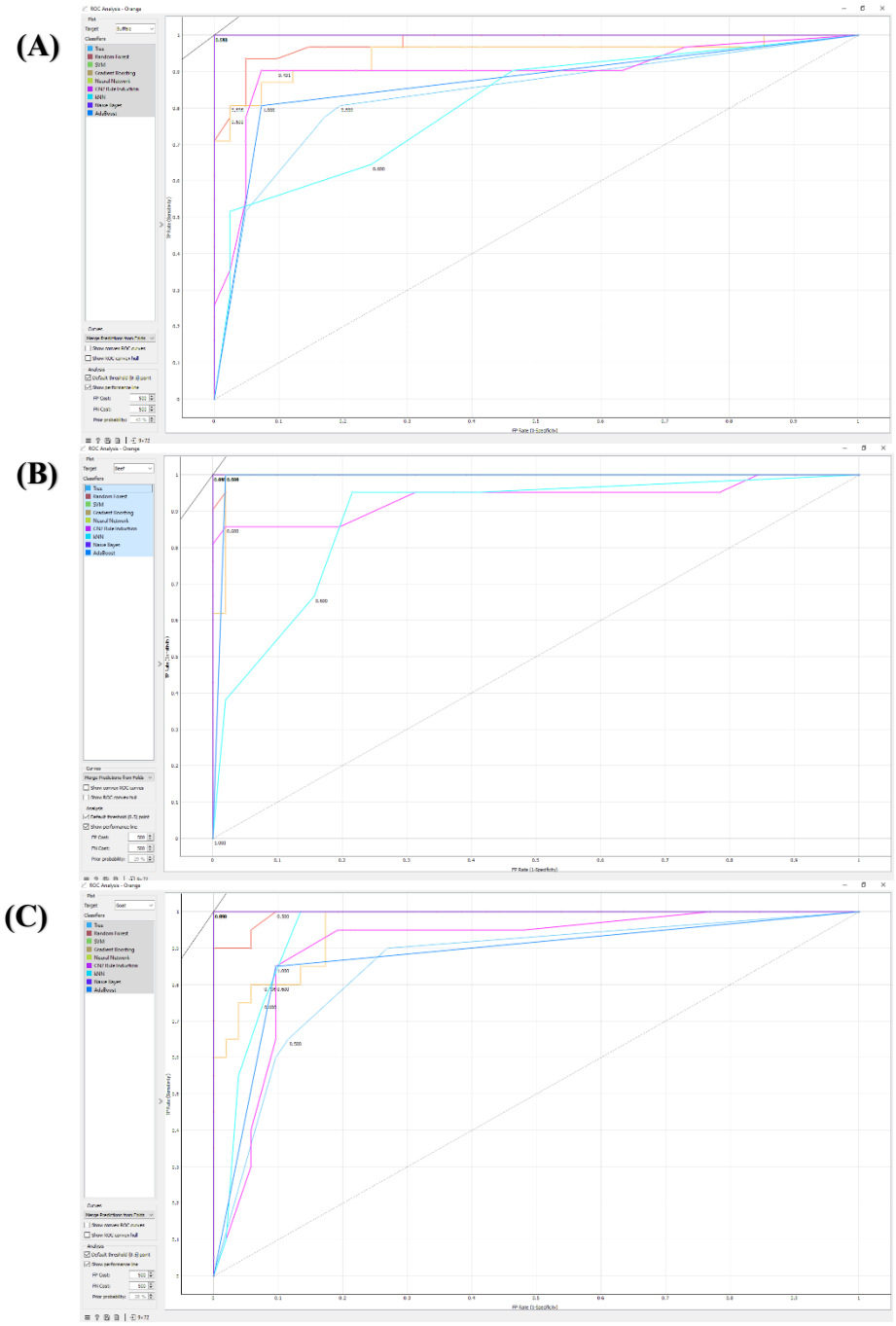

Figure 3. ROC analysis of the buffalo (A), beef (B), and goat (C) meat targets

The predicted accuracy of the classification algorithm on an unreported testing dataset is shown in Table 2. This study is delineated the models' learning transfer prioritization and prediction abilities on the newly collected data. The classifiers were evaluated by applying the same metrics used previously: AUC, CA, F1 score, Precision, Recall, and MCC. Support Vector Machine, Random Forest, Tree, Gradient Boosting, Neural Network, CN2 Rule Induction, Naïve Bayes, and AdaBoost ran without any issues on the test data. Notwithstanding, KNN was less efficient compared to other classifiers in the tested data.

Table 2. Average performance of the classification algorithms on the testing dataset

Model	AUC	CA	F1	Prec	Recall	MCC
Tree	1.000	1.000	1.000	1.000	1.000	1.000
Random Forest	1.000	0.933	0.934	0.945	0.933	0.906
Support Vector Machine	1.000	0.967	0.967	0.970	0.967	0.951
Gradient Boosting	1.000	0.900	0.900	0.925	0.900	0.864
Neural Network	1.000	1.000	1.000	1.000	1.000	1.000
CN2 Rule Induction	0.986	0.900	0.899	0.909	0.900	0.852
k-Nearest Neighbors	0.866	0.767	0.761	0.763	0.767	0.648
Naïve Bayes	1.000	1.000	1.000	1.000	1.000	1.000
AdaBoost	0.972	0.967	0.966	0.969	0.967	0.951

AUC = Area Under the Curve, CA = Classification Accuracy, F1 = F1 score, Prec = Precision, Recall = True Positive Rate, and MCC = Matthews Correlation Coefficient

Discussion

It is vital to realize that the best classifier for this task will change depending on the dataset and the desired balance between different metrics (Evidently, 2024). The Support Vector Machine, Neural Network, and Naïve Bayes accurately classified buffalo, beef, and goat meat for this investigation. However, a small dataset is prone to overfitting and will not perform well in new scenarios (Han and Jiang, 2014; Mahmud *et al*., 2020) or real-world datasets. Neural Networks have deep architecture and thus, the risk of overfitting is closely connected. It is only when kernel and regularization are applied that the Support Vector Machine context becomes complete. The performance of the Naïve Bayes algorithm can vary when the scenario is different. This study showed that it provided better outcomes than a Support Vector Machine when the misclassification cost is exceptionally high (Huesmann *et al*., 2020; Ibrahim *et al*., 2009; Tantuğ and Eryiğit, 2006). Identifying the appropriate classifier, applying regularization, and performing proper validation analysis are found to be important step to avoid overfitting and ensured that the model to run

successfully in new data (Hamidi *et al*., 2015; Ibrahim *et al*., 2009; Tagliaferri *et al*., 2015).

The experiments demonstrated that Random Forest algorithms had higher accuracy with imbalanced datasets in contrast to Gradient Boosting (Sopiyan *et al*., 2022). Moreover, Random Forest can provide significant performance benefits when combined with decision fusion (Reddy *et al*., 2022). However, Gradient Boosting impressed in instances where the priority is to minimize false negatives (Wang, 2023) and in the case of outliers (Pandurang Adi *et al*., 2020). Nevertheless, the classifiers have always been highly rated in various evaluation metrics, demonstrating their usefulness in a wide range of applications (Nainggolan and Sinaga, 2023). The low prediction effectiveness of the Tree, CN2 Rule Induction, k-Nearest Neighbors, and Ada Boost methods are found to be concerned in several factors, such as the presence of bad hubs (Buza *et al.*, 2015), sensitivity to the problem of high dimensionality (Rani, 2017), and equal attribute weighting, which can lead to poor performance (Rani, 2017). Additionally, the k-Nearest Neighbors method is prone to overfitting, computationally complex, and sensitive to the choice of k, particularly in large datasets (Onyezewe *et al.*, 2021). Finally, the AdaBoost classifier obtained an AUC of 0.909, indicating high performance. Its CA and F1 scores were also high (0.875), making it another strong contender for this task.

The development of perfect classification by Support Vector Machine, Neural Network, and Naïve Bayes revealed two facets as they are measured on the same level. First, the dataset used for training and testing was likely to have a structured form that made it possible to differentiate fine details of classes (Ioannou and Vassiliou, 2021). Moreover, excellent performance might result from fine-modal tuning and model optimization (Khorramifar *et al*., 2022). In addition, the algorithm's error-free accuracy proved its proper ability to handle the target problem (Alqasmi *et al*., 2020). In the second place, the high precision seems to be leaning toward the robustness and generalizability of these models, meaning that they can distinguish previously unseen data with the same expected result (Kumar *et al*., 2022).

When a model is tested on the test dataset and the test performance showed closely relationship with the training data performance, the model is considered to be appropriated a balance for the training data and unseen data (Hastie *et al*., 2009). The model demonstrated an impressive resulted not only discovered but also remembered the patterns existing in the original data, thereby offering highly accurate predictions. It combats accuracy and complexity, thereby preventing users from falling into the trap of overfitting the training data (Bishop, 2006). Through the test results, it can be concluded that due to its performance value, the model is reliable for use in predicting future unseen output data (Russel and

Norvig, 2010). It is noteworthy that the proposed approach is accurately obtained the data running relationship, and never overfitted the shared data, and handles the generalization of unseen data. Discovering the difference in the results is obtained by a given model on the training and test data can help to determine whether it is overfitting or underfitting.

The overfitting problem is common in ML; a model can perform excellently on training data but fails when applied to test data (Prieditis and Sapp, 2013; Trivedi *et al*., 2021). This situation is manifested when the model took the training data that the noise and outliers are captured. As a result, the model is worked well on known data but performed poorly when confronted with unseen data. Furthermore, this problem is occured when the model become too complex and began to learn noise and random variations in the dataset rather than concentrating on the underlying patterns (Bishop, 2006). Moreover, overfitting can occur when there are more features or parameters than the number of examples in training. It can also occur if the model is too complex for the specific issue being addressed (Goodfellow *et al*., 2016). According to James *et al*. (2013), this is known as the "over-estimation" of model performance. Regularization measures such as L1 and L2 can be applied to address this problem. The preferred methods were used to reduce the model's capacity and restricted it from overfitting noise in the training datasets (James *et al*., 2013).

Furthermore, if an artificial neural network meets its training data accuracy but fails its test data accuracy, then overfitting is indicated. Accordingly, underfitting problems emerge when models are exhibited poor performance, not only on training but also on test datasets, although the models may fit the training data very well (Li *et al*., 2021b). The algorithm may be under duress when identifying fundamental patterns in the training data, leading to poor performance on the training set (Hastie *et al*., 2009). To avoid the underfitting problem, the use of different strategies like regularization and feature selection, should be employed. The techniques that could be attempted include the model complexity adjustment, using dropout as a regularization technique during the early phases of training, and generating special approaches for domain-specific areas, respectively. Weakening results on the training set and failure to replicate the same good outcome by the model on the test set is typically demonstrated underfitting.

It can be concluded that the Support Vector Machine, Neural Network, and Naïve Bayes, were shown to be leading three models in properly distinguishing meat quality from buffalo, beef, and goat meat. In future studies, it would undoubtedly be helpful to investigate the applicability of these approaches to various other meat categories. The effects of different preprocessing information strategies on model precision should be analysed. Furthermore, the development

of the models using large datasets or a variety of information on meat quality would substantially improve the model's accuracy for real-world application.

References

Alqasmi, D. U., Alzuhair, A. and Bama'bad, A. (2020). Enhanced system for computer-aided detection of mri brain tumors. Signal & Image Processing: An International Journal, 11:25-31.

AMSA (2012). Meat Color Measurement Guidelines. American Meat Science Association, IL, 124 p. Retrieved from: https://meatscience.org/publications-resources/printed-publications/amsa-meat-color-measurement-guidelines

AMSA (2016). Research Guidelines for Cookery, Sensory Evaluation, and Instrumental Tenderness Measurements of Meat, 2nd ed. American Meat Science Association, IL, 105 p. Retrieved from: https://meatscience.org/publications-resources/printed-publications/sensory-and-tenderness-evaluation-guidelines

Arora, S. (2018). Mathematics of machine learning: An introduction. Proceedings of the International Congress of Mathematicians, 1:377-390.

Association of Official Analytical Chemists (AOAC). (2012). Official Method of Analysis (Association of Official Analytical Chemist (ed.); 19th ed.).

Ayaz, H., Ahmad, M., Sohaib, A., Yasir, M. N., Zaidan, M. A., Ali, M., Khan, M. H. and Saleem, Z. (2020). Myoglobin-based classification of minced meat using hyperspectral imaging. Applied Sciences, 10:6862.

Bishop, C. M. (2006). Pattern Recognition and Machine Learning. Springer New York.

Bourne, M. C. (1978). Texture profile analysis. Food Technology, 32:62-67.

Buza, K., Nanopoulos, A. and Nagy, G. (2015). Nearest neighbor regression in the presence of bad hubs. Knowledge-Based Systems, 86:250-260.

Dayal, M., Gupta, M., Gupta, M., Bara, A. R. and Chaubey, C. (2023). Introduction to machine learning methods with application in agriculture. In Applying Drone Technologies and Robotics for Agricultural Sustainability. Retrieved from: https://doi.org/10.4018/978-1-6684-6413-7.ch012

Demšar, J., Curk, T., Erjavec, A., Gorup, Č., Hočevar, T., Milutinovič, M., Možina, M., Polajnar, M., Toplak, M., Starič, A., Štajdohar, M., Umek, L., Žagar, L., Žbontar, J., Žitnik, M. and Zupan, B. (2013). Orange: Data Mining Toolbox in Python. Journal of Machine Learning Research, 14:2349-2353.

Domingos, P. (2012). A few useful things to know about machine learning. Communications of the ACM, 55:78-87.

Duarte, D. and Ståhl, N. (2019). Machine learning: A concise overview. Studies in Big Data, 46:27-58.

Elsaadouny, M., Barowski, J. and Rolfes, I. (2021). Unsupervised learning implementation for SAR Images Clustering. International Conference on Electromagnetics in Advanced Applications, ICEAA 2021, 104. Retrieved from: https://doi.org/10.1109/ICEAA52647.2021.9539661

Evidently, A. I. (2024). A complete guide to classification metrics in machine learning. Retrieved from: https://www.evidentlyai.com/classification-metrics

Forsyth, D. (2018). Probability and Statistics for Computer Science. In: Probability and Statistics for Computer Science. Springer International Publishing.

Goodfellow, I., Bengio, Y. and Courville, A. (2016). Deep Learning. MIT Press.

Gupta, V., Mishra, V. K., Singhal, P. and Kumar, A. (2022). An overview of supervised machine learning algorithm. Proceedings of the 2022 11th International Conference on System Modeling and Advancement in Research Trends, SMART 2022, pp.87-92.

Hamidi, O., Poorolajal, J., Sadeghifar, M., Abbasi, H., Maryanaji, Z., Faridi, H. R. and Tapak, L. (2015). A comparative study of support vector machines and artificial neural networks for predicting precipitation in Iran. Theoretical and Applied Climatology, 119:723-731.

Han, H, and Jiang, X. (2014). Overcome Support Vector Machine diagnosis overfitting. Cancer Informatics, 13:145-158.

Hastie, T., Tibshirani, R., and Friedman, J. (2009). The Elements of Statistical Learning. Springer New York.

Head, M. and Aloqaily, W. (2022). 9 - Machine learning approach for seismic assessment. In: Seismic Evaluation, Damage, and Mitigation in Structures. Woodhead Publishing, pp.237-247.

Honikel, K. O. (1998). Reference methods for the assessment of physical characteristics of meat. Meat Science, 49:447-457.

Huesmann, K., Klemm, S., Linsen, L. and Risse, B. (2020). Exploiting the full capacity of deep neural networks while avoiding overfitting by targeted sparsity regularization. Retrieved from: https://doi.org/10.48550/arxiv.2002.09237

Ibrahim, Z., Isa, D., Rajkumar, R. and Kendall, G. (2009). Document zone content classification for technical document images using Artificial Neural Networks and Support Vector Machines. 2nd International Conference on the Applications of Digital Information and Web Technologies, 345-350. Retrieved from: https://doi.org/10.1109/ICADIWT.2009.5273957

Ioannou, C. and Vassiliou, V. (2021). Network attack classification in IOT using support vector machines. Journal of Sensor and Actuator Networks, 10:58.

James, G., Witten, D., Hastie, T. and Tibshirani, R. (2013). An Introduction to Statistical Learning: with Applications in R. Springer.

Jordan, M. I., and Mitchell, T. M. (2015). Machine learning: Trends, perspectives, and prospects. Science, 349:255-260.

Khorramifar, A., Rasekh, M., Karami, H., Covington, J. A., Derakhshani, S. M., Ramos, J. and Gancarz, M. (2022). Application of MOS gas sensors coupled with chemometrics methods to predict the amount of sugar and carbohydrates in potatoes. Molecules, 27: 3508.

Kumar, H. and Hasija, Y. (2021). Machine Learning in Medical Image Processing. Smart Innovation, Systems and Technologies, 195:377-383.

Kumar, R., Aggarwal, Y. and Kumar Nigam, V. (2022). Heart rate dynamics in the prediction of coronary artery disease and myocardial infarction using artificial neural network and support vector machine. Journal of Applied Biomedicine, 20:70-79.

Li, A., Li, C., Gao, M., Yang, S., Liu, R., Chen, W. and Xu, K. (2021a). Beef cut classification using multispectral imaging and machine learning method. Frontiers in Nutrition, 8: 755007.

Li, Q., Yan, M. and Xu, J. (2021b). Optimizing convolutional neural network performance by mitigating underfitting and overfitting. 19th International Conference on Computer and Information Science (ICIS), pp. 126-131.

Mahmud, M. S., Huang, J. Z. and Fu, X. (2020). Variational autoencoder-based dimensionality reduction for high-dimensional small-sample data classification. International Journal of Computational Intelligence and Applications, 19:2050002.

Malikhah, M., Sarno, R. and Sabilla, S. (2021). Ensemble learning for optimizing classification of pork adulteration in beef based on electronic nose dataset. International Journal of Intelligent Engineering and Systems, 14:44-55.

Medeiros, E. C., Almeida, L. M. and Filho, J. G. de A. T. (2021). Computer vision and machine learning for tuna and salmon meat classification. Informatics, 8:70.

Nainggolan, S. P. and Sinaga, A. (2023). Comparative analysis of accuracy of random forest and gradient boosting classifier algorithm for diabetes classification. Sebatik, 27:97-102.

Onyezewe, A., Kana, A. F., Abdullahi, F. B. and Abdulsalami, A. O. (2021). An enhanced adaptive k-Nearest Neighbor classifier using simulated annealing. International Journal of Intelligent Systems and Applications, 13:34-44.

Pandurang Adi, S., Bettadapura Adishesha, V., Vaidyanathan Bharadwaj, K. and Narayan, A. (2020). Earthquake damage prediction using random forest and gradient boosting classifier. American Journal of Biological and Environmental Statistics, 6:58.

Peng, G. C. Y., Alber, M., Tepole, A. B., Cannon, W. R., De, S., Dura-Bernal, S., Garikipati, K., Karniadakis, G., Lytton, W. W., Perdikaris, P., Petzold, L. and Kuhl, E. (2021). Multiscale modeling meets machine learning: What can we learn? Archives of Computational Methods in Engineering: State of the Art Reviews, 28:1017-1037.

Phoemchalard, C., Tathong, T., Pornanek, P., Uriyapongson, S. and Cherdthong, A. (2021). Quality attributes of buffalo meat using precision sous vide cooking device. ARPN Journal of Engineering and Applied Sciences, 16:2117-2125.

Phoemchalard, C., Uriyapongson, S., Tathong, T. and Pornanek, P. (2022). ^{1}H NMR metabolic profiling and meat quality in three beef cattle breeds from northeastern Thailand. Foods, 11:3821.

Prieditis, A. and Sapp, S. (2013). Lazy Overfitting Control. In P. Perner (Ed.), Machine Learning and Data Mining in Pattern Recognition. Springer Berlin Heidelberg, pp.481-491.

Qiu, J., Lin, Y., Wu, J., Xiao, Y., Cai, H. and Ni, H. (2024). Rapid beef quality detection using spectra pre-processing methods in electrical impedance spectroscopy and machine learning. International Journal of Food Science and Technology, 59:1624-1634.

Rani, P. (2017). A Review of various KNN techniques. International Journal for Research in Applied Science and Engineering Technology, 5:1174-1179.

Rashidi, H. H., Tran, N. K., Betts, E. V., Howell, L. P. and Green, R. (2019). Artificial Intelligence and Machine Learning in Pathology: The Present Landscape of Supervised Methods. Academic Pathology, 6:2374289519873088.

Reddy, M. N., Latchmana Kumar, M. R., Kumar, P. B. S., Thirumalai, S. and Nirmala Devi, M. (2022). Performance enhancement by tuning hyperparameters of random forest classifier for hardware trojan detection. In M. Saraswat, S. Roy, C. Chowdhury, and A. H. Gandomi (Eds.), Proceedings of International Conference on Data Science and Applications, pp.177-191.

Russell, S. and Norvig, P. (2010). Artificial Intelligence: A Modern Approach (3rd ed.). Prentice Hall.

Sathya, D., Sudha, V. and Jagadeesan, D. (2022). Application of machine learning techniques in healthcare. In: Research Anthology on Machine Learning Techniques, Methods, and Applications. https://doi.org/10.4018/978-1-6684-6291-1.ch067

Sen, R. and Das, S. (2023). Unsupervised Learning. Indian Statistical Institute Series, 305-318. https://doi.org/10.1007/978-981-19-2008-0_21

Setiadi, I. C., Hatta, A. M., Koentjoro, S., Stendafity, S., Azizah, N. N. and Wijaya, W. Y. (2022). Adulteration detection in minced beef using low-cost color imaging system coupled with deep neural network. Frontiers in Sustainable Food Systems, 6:1073969

Sharma, R. and Bharti, V. (2021). A Review Towards Machine Learning Spectrum for Bioinformatics. Proceedings - 2021 3rd International Conference on Advances in Computing, Communication Control and Networking, ICAC3N 2021, 199-202.
Sopiyan, M., Fauziah, F. and Wijaya, Y. F. (2022). Fraud detection using random forest classifier, logistic regression, and gradient boosting classifier algorithms on credit cards. JUITA: Jurnal Informatika, 10:77.
Tagliaferri, F., Viola, I. M. and Flay, R. G. J. (2015). Wind direction forecasting with artificial neural networks and support vector machines. Ocean Engineering, 97:65-73.
Tantuğ, A. C. and Eryiğit, G. (2006). Performance analysis of naïve bayes classification, support vector machines and neural networks for spam categorization. In Applied Soft Computing Technologies: The Challenge of Complexity, Springer-Verlag, pp.495-504.
Thakur, U. K. (2023). The role of machine learning in customer experience. In: Handbook of Research on AI and Machine Learning Applications in Customer Support and Analytics. Retrieved from: https://doi.org/10.4018/978-1-6684-7105-0.ch005
Torres, J. P., Caro, A., Avila, M. D. M., Perez-Palacios, T., Antequera, T. and Rodriguez, P. G. (2022). A computer-aided inspection system to predict quality characteristics in food technology. IEEE Access, 10:71496-71507.
Trivedi, U. B., Bhatt, M. and Srivastava, P. (2021). Prevent Overfitting Problem in Machine Learning: A Case Focus on Linear Regression and Logistics Regression. In P. K. Singh, Z. Polkowski, S. Tanwar, S. K. Pandey, G. Matei, and D. Pirvu (Eds.), Innovations in Information and Communication Technologies (IICT-2020). Springer International Publishing, pp.345-349.
Wang, Y. (2023). Comparison of machine learning strategies in hazardous asteroids prediction. Highlights in Science, Engineering and Technology, 39:201-208.
Yang, F., Sun, J., Cheng, J., Fu, L., Wang, S. and Xu, M. (2023). Detection of starch in minced chicken meat based on hyperspectral imaging technique and transfer learning. Journal of Food Process Engineering, 46:e14304.

(Received: 10 August 2024, Revised: 18 November 2024, Accepted: 19 November 2024)