Strengthening of the Italian Research Infrastructure for Metrology and Open Access Data in support to the Agrifood (METROFOOD-IT)

Michele Magarelli[1]

[1]Department of Soil, Plant and Food Sciences (DISSPA), University of Bari Aldo Moro (UNIBA), Via Giovanni Amendola, 165/a, 70126 Bari (BA)

Abstract

This research explores video data explanations, incorporating temporal information, via XAI methods to enhance reliability for surgical training. It aims to detect learning biases in DNNs used for surgical skill assessment. The broader objective is to evaluate XAI performance with complex models and real-world data. This document details the context, methodology, preliminary findings, and future contributions.

1. Context and Motivation

In recent years, the demand for high-quality and region-specific food products has significantly increased, driven by growing consumer awareness and increasingly stringent regulations on food safety and traceability. Ensuring the authenticity of agri-food products has become essential to combat food fraud, protect geographical indications, and promote transparency throughout the entire production chain. Traceability not only enhances the value of local specialties and quality certifications (such as PDO and PGI) but also plays a crucial role in detecting potential contamination or adulteration that could compromise consumer safety[1]. Currently, food origin determination relies on various methodologies, including certification systems, blockchain technology, spectroscopic and physicochemical analyses[2]. However, these techniques have limitations in terms of cost, scalability, and their ability to provide in-depth insights into all aspects related to production and quality. In this context, Artificial Intelligence (AI)—particularly Machine Learning (ML)—is emerging as an innovative tool capable of analyzing large datasets and identifying complex relationships between the chemical, physical, and environmental characteristics of food products. By applying ML models, it is possible to develop advanced systems for uncovering hidden patterns in data, improving accuracy in food origin determination, and optimizing quality control processes. These models can be used to classify products based on specific parameters, predict quality variations depending on production conditions, and detect anomalies that may indicate fraud or contamination. However, one of the main challenges in applying AI to this field is the lack of transparency in ML-based decision-making. Many AI techniques, especially advanced models such as deep neural networks, function as “black boxes,” making it difficult to understand how and why a certain prediction is generated. To address this challenge, this study—conducted within the framework of the METROFOOD-IT project—adopts an Explainable Artificial Intelligence (XAI) approach [3]. In this context the project is designed to support research and innovation in the agrifood sector by providing integrated services that enhance digitalization and improve the efficiency, traceability, and sustainability of agrifood systems. A key objective is to increase the reliability of products and processes while ensuring that citizens, authorities, and food system stakeholders have access to transparent and trustworthy information. The project will also define a structured framework for service provision, enabling both transnational and virtual access to advanced research services that support interdisciplinary studies in the agrifood sector. XAI provides methods to interpret and explain AI-generated results, allowing

Late-breaking work, Demos and Doctoral Consortium, colocated with The 3rd World Conference on eXplainable Artificial Intelligence: July 09–11, 2025, Istanbul, Turkey
michele.magarelli@uniba.it (M. Magarelli)
0009-0008-9119-5638 (M. Magarelli)

experts to identify the key factors influencing food quality and traceability[4]. The goal is to improve the transparency of ML-based analyses, offering reliable tools for food industry professionals to assess and monitor product quality. This advanced approach applies to various food categories, including olive oil, tomatoes, and mozzarella, enabling a more comprehensive traceability system and enhancing certification processes. Moreover, the integration of Machine Learning and XAI contributes to strengthening trust among stakeholders in the food sector, providing a clearer and more detailed understanding of the mechanisms determining food authenticity and quality. In an increasingly demanding and regulated global market, ensuring food quality, safety, and transparency is now more crucial than ever.

2. Key Related Work

The integration of Machine Learning (ML) and Explainable Artificial Intelligence (XAI) in food quality assessment has gained increasing attention in recent years [5]. Several studies have explored ML techniques for detecting food adulteration, assessing quality parameters, and improving traceability, while XAI has been proposed as a solution to enhance model interpretability and stakeholder trust. This section reviews the most relevant contributions in this domain, focusing on food authenticity verification, spectral data analysis, and XAI applications in food science. Various machine learning approaches have been applied to assess food quality, particularly using spectral techniques such as Near-Infrared (NIR) and Infrared (IR) spectroscopy. Studies have demonstrated that deep learning models, including convolutional neural networks (CNNs) and ensemble methods, can effectively classify food products based on their chemical and spectral fingerprints. For example, research on dairy products, including mozzarella, has highlighted the potential of ML for distinguishing authentic samples from fraudulent or lower-quality variants, often leveraging chemometric techniques in combination with supervised learning models [6]. While ML models have shown high accuracy in food classification tasks, their adoption is often limited by the lack of interpretability. Recent works have introduced XAI techniques, such as SHAP and LIME, to provide insight into feature importance and model decisions in food science applications [7]. Studies in this area have demonstrated that XAI can help identify key chemical markers responsible for quality classification, improving trust and acceptance of AI-driven food assessment tools among scientists, regulators, and industry stakeholders.

3. Research Framework

Food quality and traceability are critical challenges in modern food systems, requiring robust analytical methods to ensure safety, authenticity, and compliance with regulations. The advent of Machine Learning (ML) and Explainable Artificial Intelligence (XAI) provides new opportunities to enhance the efficiency, accuracy, and interpretability of food quality assessments. METROFOOD framework, aims to establish a research infrastructure for metrology and traceability in food quality and safety.

3.1. Hypothesis

The hypotheses guiding this research are formulated to address key challenges in food quality assessment, authenticity verification, and traceability within the METROFOOD framework. These hypotheses center on the effectiveness of ML and XAI techniques in improving the accuracy, transparency, and reliability of food quality models. Each hypothesis is designed to explore the potential of these methodologies to revolutionize food safety and quality control by providing more precise, interpretable, and trustworthy analytical tools. The first hypothesis suggests that integrating ML and XAI will enhance the accuracy and interpretability of food quality models, making them more applicable in real-world scenarios. ML models, especially deep learning, are powerful in detecting patterns within complex datasets like spectral and chemical data. However, their lack of transparency makes them harder to trust. By combining these models with XAI techniques, the decision-making process becomes more understandable, allowing stakeholders to confidently apply these models within METROFOOD [8]. The second hypothesis emphasizes that XAI enhances the transparency of ML models, which is crucial for building trust among food scientists, regulators, and consumers. Even though ML can provide accurate predictions, it often lacks interpretability. XAI methods can explain the factors influencing these predictions, making the results more accessible and reliable for stakeholders who need to understand how food quality decisions are made[9]. The third hypothesis posits that advanced ML techniques, such as deep learning and ensemble models, can process complex food-related datasets, like spectral and chemical data, to identify quality indicators and detect anomalies. These models are well-suited for extracting valuable insights from large, intricate datasets, improving the reliability of food quality assessments and helping to detect issues like fraud or contamination. Finally, the fourth hypothesis proposes that the integration of XAI-driven ML models in METROFOOD will lead to more reliable food quality control, improving traceability and ensuring compliance with safety standards. By providing interpretable, real-time insights, these models can support food safety monitoring, making decision-making more efficient and trustworthy, while enhancing transparency and accountability.

3.2. Objectives

The primary objective of this research is to develop and validate Machine Learning (ML) and Explainable Artificial Intelligence (XAI) approaches for food quality assessment within the METROFOOD initiative [10]. This involves designing and evaluating ML models capable of analyzing food-related data to assess quality, authenticity, and traceability. The study will explore different ML techniques, such as Random Forest, Support Vector Machines, and Neural Networks, comparing their performance based on key metrics like accuracy, precision, recall, and F1-score. A crucial aspect of this research is investigating the role of XAI in improving the interpretability and transparency of ML models in food science applications. Various explainability techniques, including SHAP, LIME, and Attention Mechanisms, will be implemented to make ML-based food quality assessments more accessible and understandable for domain experts. The effectiveness of these XAI methods will be evaluated by assessing their alignment with expert knowledge and their ability to enhance trust in AI-driven decision-making. The study will also apply ML and XAI methodologies to real-world datasets from METROFOOD, focusing on spectral analysis data (e.g., NIR, IR) and microbiome composition[11] profiles to distinguish high-quality food products from substandard or fraudulent ones. The potential of ML to detect anomalies and deviations in food production and supply chains will be explored, contributing to improved traceability and quality assurance. Finally, the research aims to enhance data-driven decision-making processes in food quality control by integrating ML and XAI methodologies into existing food analysis frameworks. A key goal is to develop an XAI-based framework that supports researchers, regulatory bodies, and industry professionals in understanding and interpreting ML-generated insights. Additionally, the study will examine potential deployment strategies for integrating ML models into METROFOOD's digital infrastructure, ensuring their practical applicability in real-world food quality monitoring and assessment.

4. Research Approach

Our research approach is based on a quantitative methodology aimed at analyzing and predicting food quality using advanced Machine Learning (ML) and Explainable AI (XAI) techniques[12]. The study focuses on analyzing food data from various sources, such as sensors, laboratories, and existing databases, in order to develop predictive models for tracking the quality of products such as olive oil, tomatoes, and mozzarella. These data will be used to train ML models with the goal of identifying and predicting the quality and authenticity of each product. The approach will primarily rely on classification and regression techniques, to address the complexity of the data and the possible variables influencing food quality. To enhance the reliability and interpretability of the results, Explainable AI (XAI) methods will also be employed to make the predictive models not only effective but also understandable to stakeholders, such as producers and consumers, ensuring transparency in automated decisions.

4.1. Materials and Methods

The project involves the use of data acquired from various partners within the METROFOOD consortium to ensure a comprehensive analysis of food quality and traceability. However, at present, the only published study is based on data related to Mozzarella di Bufala Campana DOP. Mozzarella di Bufala Campana is a soft, fresh, stretched-curd cheese traditionally produced in the provinces of Caserta and Salerno (Italy). Production also takes place in selected localities of the metropolitan city of Naples, as well as in southern Lazio, northern Apulia, and the municipality of Venafro in Molise. Mozzarella di Bufala Campana is often known as "white gold" in homage to the cheese's prized nutritional and taste qualities. It was granted Protected Designation of Origin (PDO) status in 1996. Protected Designation of Origin (PDO) is a certification that guarantees the authenticity and quality of food products linked to specific geographical regions. The data used in this study, described in Table 1, were obtained from the microbiological analysis of the microbiome of 65 samples of Mozzarella di Bufala Campana DOP, collected from 30 dairies in the province of Salerno and 35 dairies in the province of Caserta [13].

n samples from Salerno	30
n samples from Caserta	35
type of input variables	microbiome relative abundance
n input variables for each sample	139

Table 1
Description of samples and input variables

These samples underwent thorough examination in the laboratories of the Microbiology Division of the Department of Agricultural Sciences at the University of Naples Federico II [14]. All the dairies selected for this study are located within the production area defined by the official production regulations and exclusively produce Mozzarella di Bufala Campana DOP.

The collected data were analyzed using Machine Learning (ML) algorithms for both classification and regression tasks. For classification, models such as EXtreme Gradient Boosting (XGBoost)[15] and Random Forest[16], and neural networks were employed to categorize the data based on predefined labels. For regression tasks, linear regression, decision tree-based models, and deep learning approaches were utilized to predict continuous outcomes. To ensure the reliability and robustness of the analysis, model performance was evaluated using appropriate metrics for each task. Classification models were assessed based on accuracy, precision, recall, F1-score and AUC (Area Under the Curve), while regression models were evaluated using Mean Absolute Error (MAE), root mean squared error (RMSE), and R-squared (R^2). To optimize performance and prevent overfitting, hyperparameter tuning was conducted using grid search and random search techniques, and model validation was performed through k-fold cross-validation[17]. Additionally, feature selection and engineering techniques were applied to enhance model interpretability and predictive accuracy. Dimensionality reduction methods, such as Principal Component Analysis (PCA), were used when necessary to manage high-dimensional data. All analyses were conducted using Python-based frameworks, including Scikit-learn, TensorFlow, and PyTorch, ensuring reproducibility and scalability of the results. To further enhance model interpretability and transparency, Explainable Artificial Intelligence (XAI) techniques were integrated into the analysis[18]. Among the various XAI methods, SHapley Additive exPlanations (SHAP) and feature importance analysis were employed to provide deeper insights into the decision-making process of the models. SHAP offers a systematic approach to explaining model predictions by assigning an importance value to each feature based on its contribution to the final output. By quantifying the individual impact of each feature, SHAP values enable a clearer understanding of how different variables influence model predictions. The SHAP value for a given feature is computed by averaging its marginal contribution across all possible feature subsets, ensuring a fair and comprehensive attribution of importance. This methodology provides a robust framework for interpreting complex ML models, particularly when dealing with high-dimensional datasets. In addition to SHAP, feature importance analysis was used to rank variables based on their influence on model performance. This technique is particularly relevant for tree-based models, where built-in feature importance metrics allow for an intuitive assessment of the most critical predictors. Understanding feature importance not only aids in model refinement but also provides valuable insights for domain experts, facilitating more informed decision-making.

4.2. Rationale for Testing the Research Hypothesis

The research hypotheses focus on the effectiveness of Machine Learning methods in tracking food quality and the importance of interpretability through XAI techniques [19]. The primary hypothesis posits that the use of Machine Learning algorithms to analyze food data can significantly improve accuracy in monitoring and predicting food quality compared to traditional methods. Machine Learning models are particularly well-suited to handle large volumes and complexities of data, as they are capable of identifying hidden patterns and making accurate predictions. The secondary hypothesis suggests that integrating XAI techniques will improve the transparency and reliability of the models, making the results more understandable and increasing user trust in food traceability systems. Testing these hypotheses is crucial to determine whether the adoption of advanced Machine Learning technologies can indeed optimize the traceability process and improve the overall quality of the food system. In particular, the proposed research approach will provide concrete answers to how XAI can be used to make Machine Learning solutions more feasible/acceptable in both industrial and regulatory contexts.

5. Preliminary Results

As part of the ongoing research, preliminary analyses have been conducted to explore the potential of ML and XAI in food quality assessment, with a specific focus on the classification and characterization of mozzarella samples [20]. Initial results suggest that ML models can effectively differentiate between DOP and noDOP samples based on the microbiome composition. However, challenges have emerged regarding data variability, model generalization, and the interpretability of certain predictive features. This section discusses these preliminary findings and their implications for refining the models and improving their robustness within real-world food quality control frameworks. This study involved evaluating the effectiveness of three supervised machine learning algorithms, namely XGBoost, Random Forest, and a complex Multi-Layer Perceptron network. The key steps of the analysis are depicted in the flowcharts provided in the Figure 1.

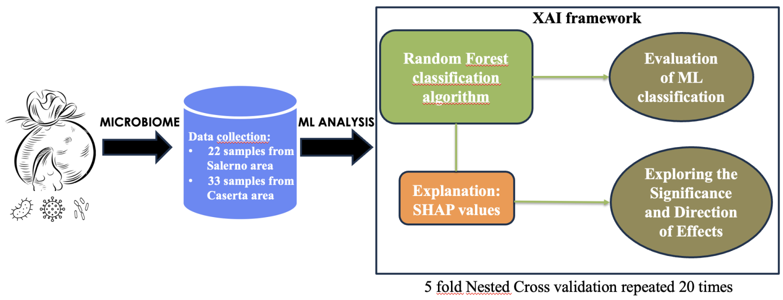

Figure 1: The diagram delineates the sequence of the conducted analysis. The dataset comprised 55 samples, 22 from Salerno and 33 from Caserta provinces.This dataset served to evaluate a 5-fold cross-validation, with subsequent analysis of XGB Classifier. Finally, the evaluation metrics used to assess the performance of RF classifier.

The results have been obtained following a 5-fold repeated 20 times cross-validation procedure on the validation set. This methodology allows us to assess the effectiveness of our algorithm in a robust and reliable manner.The analysis revealed that the Random Forest classifier outperformed the others, demonstrating the highest Area Under the Curve (AUC) value of 0.93 ± 0.10 and the top accuracy score of 0.87 ± 0.11. Table 2 provides a comprehensive comparison of the three models based on their AUC and accuracy scores.

Classifier	Accuracy	AUC
XGB	0.82 ± 0.12	0.87 ± 0.11
RF	0.87 ± 0.11	0.93 ± 0.10
MLP	0.68 ± 0.13	0.78 ± 0.11

Table 2
Comparison between evaluation metrics of XGBoost (XGB), Random Forest (RF), and Multi-Layer Perceptron (MLP) classifiers.

After conducting cross-validation, the outcomes were then utilized to compute feature importance employing SHAP. Through a SHAP analysis, the 20 most important feature were identified, deriving from the analysis of the microbiota 65 samples. In the SHAP plot in figure 2 it is evident how certain features, such as Lactococcus lactis and Moraxella osloensis, contribute significantly to the model's prediction. The feature Lactobacillus helveticus is important for the model's interpretability, as the colored points are well distinguished, and red points indicate that high values of that bacterium have influenced Salerno class, and vice versa. This suggests that these elements play a crucial role in the geographical discrimination of the samples.

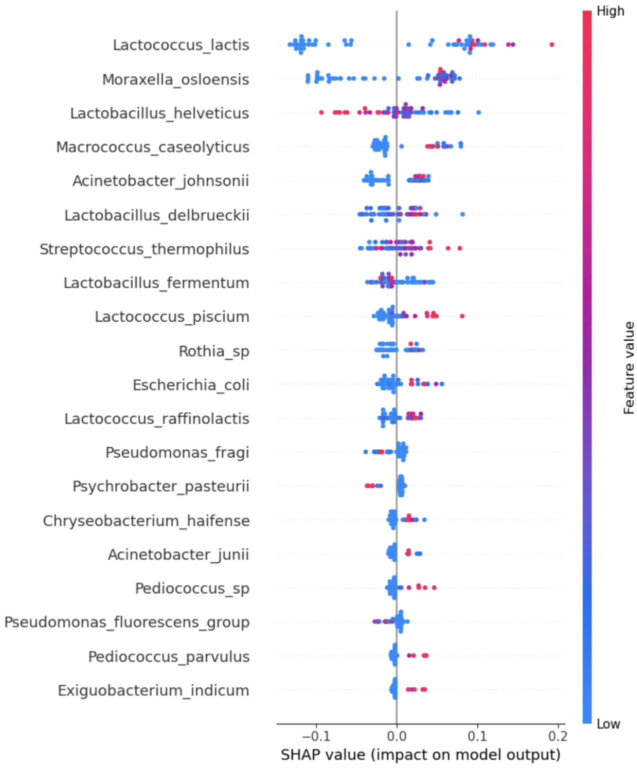

Figure 2: The SHAP (SHapley Additive exPlanations) summary plot provides an overview of the importance of features in contributing to model predictions. In this type of plot, each point represents a data instance, and the horizontal position of the point indicates how much the effect of a specific feature contributes to the change in prediction compared to the model's average prediction. The color of the point represents the value of the feature, with darker colors indicating higher values.

6. Expected Next Research Steps

The forthcoming steps in our research will focus on the advanced analysis of NIR (Near Infrared) and IR (Infrared) spectra to assess the quality and authenticity of olive oil and tomatoes. For olive oil, the primary objective will be to distinguish extra virgin olive oil (EVOO) from non-EVOO based on the ethyl esters (EE) content. Ethyl esters are key indicators of oil quality and authenticity, often linked to oxidation and the presence of adulterants. NIR and IR spectroscopy has already been employed to collect spectral data from a range of olive oil samples, both EVOO and non-EVOO. This method has provided valuable insights into the chemical composition of the oils, capturing critical quality markers such as fatty acid profile, moisture content, and oxidation levels. The next phase of the research will involve extracting relevant spectral features that correspond to ethyl esters, followed by the development of Machine Learning models, particularly classification algorithms, to differentiate EVOO from non-EVOO based on the spectral data. These models are currently being fine-tuned, and their performance will be validated through cross-validation techniques and comparisons with traditional chemical analysis methods used for determining ethyl ester content. The results of this phase are already being compiled into a manuscript for publication. Tomatoes analysis focuses on distinguishing Protected Designation of Origin (PDO) tomatoes from non-PDO varieties by analyzing the chemical characteristics of the pulp. NIR and IR spectroscopy has also been used to evaluate key chemical compounds, such as sugars, acids, and phenolic compounds, that influence the quality and authenticity of tomatoes, particularly those with PDO certification. The spectral data has already been analyzed, revealing distinct patterns related to the chemical signatures of PDO tomatoes, which are known for their specific growing conditions and quality attributes. Data has been collected from tomatoes at various stages of ripeness and from different varieties, creating a comprehensive dataset that reflects the diversity of chemical compositions. The next steps will involve further processing of the spectral data to extract relevant features, which will be used to build predictive models capable of classifying tomatoes as PDO or non-PDO based on their chemical characteristics. Calibration and validation of these models are ongoing, and comparisons with traditional quality assessments, such as sensory analysis and chemical analysis, will be carried out to ensure reliability and accuracy. Both research areas aim to develop efficient, non-invasive methods for ensuring the traceability and authenticity of olive oil and tomatoes. By integrating Explainable AI (XAI) techniques into these models, the research will ensure that the results are interpretable, allowing stakeholders, including producers and consumers, to better understand the reasoning behind the model's predictions. This approach will not only enhance the accuracy and reliability of the classification but also contribute to building trust in food quality control systems.

Finally, the research will remain open to the incorporation of new data from various sources, including industry collaborators or external datasets. This flexibility will allow for the refinement of the models and ensure their robustness, adaptability, and applicability across different types of food products. By integrating new data, the study aims to improve the accuracy and generalizability of the predictive models, thereby advancing the development of effective and universally applicable food traceability and quality control systems. This approach underscores the dynamic nature of the research, which seeks to remain responsive to emerging trends and data in the food sector.

Declaration on Generative AI

The author has not employed any Generative AI tools.

References

[1] R. Badia-Melis, P. Mishra, L. Ruiz-García, Food traceability: New trends and recent advances. a review, Food control 57 (2015) 393–401.

[2] J. Duan, C. Zhang, Y. Gong, S. Brown, Z. Li, A content-analysis based literature review in blockchain adoption within food supply chain, International journal of environmental research and public health 17 (2020) 1784.

[3] P. Di Bitonto, M. Magarelli, P. Novielli, D. Romano, D. Diacono, L. de Trizio, A. Mariano, C. Zoani, R. Ferrero, A. Manzin, et al., From data to nutrition: the impact of computing infrastructure and artificial intelligence, Exploration of Foods and Foodomics 2 (2024) 810–829.

[4] A. Corallo, M. E. Latino, M. Menegoli, F. Striani, The awareness assessment of the italian agri-food

industry regarding food traceability systems, Trends in Food Science & Technology 101 (2020) 28–37.
[5] S. Kuppusamy, M. Meivelu, L. Praburaman, M. M. Alam, A. G. Al-Sehemi, K. Anbarasu, Integrating ai in food contaminant analysis: Enhancing quality and environmental protection, Journal of Hazardous Materials Advances 16 (2024) 100509.
[6] R. J. Martin, R. Mittal, V. Malik, F. Jeribi, S. T. Siddiqui, M. A. Hossain, S. Swapna, Xai-powered smart agriculture framework for enhancing food productivity and sustainability, IEEE Access (2024).
[7] O. Buyuktepe, C. Catal, G. Kar, Y. Bouzembrak, H. Marvin, A. Gavai, Food fraud detection using explainable artificial intelligence, Expert Systems 42 (2025) e13387.
[8] P. Di Bitonto, L. De Trizio, M. Magarelli, D. Diacono, P. Novielli, D. Romano, C. Zoani, R. Bellotti, S. Tangaro, Metrofood-it: A data platform proposal using agrifood smart data model, in: 2024 IEEE International Workshop on Metrology for Industry 4.0 & IoT (MetroInd4. 0 & IoT), IEEE, 2024, pp. 117–122.
[9] A. B. Arrieta, N. Díaz-Rodríguez, J. Del Ser, A. Bennetot, S. Tabik, A. Barbado, S. García, S. Gil-López, D. Molina, R. Benjamins, et al., Explainable artificial intelligence (xai): Concepts, taxonomies, opportunities and challenges toward responsible ai, Information fusion 58 (2020) 82–115.
[10] P. Di Bitonto, L. De Trizio, M. Magarelli, D. Romano, P. Novielli, R. Ahsen, C. Manetti, C. Zoani, G. Sanesi, M. De Angelis, et al., Innovative virtual reality based training in the agri-food sector: Insights from metrofood-it, in: 2024 IEEE International Humanitarian Technologies Conference (IHTC), IEEE, 2024, pp. 1–6.
[11] A. Guidone, T. Zotta, A. Matera, A. Ricciardi, F. De Filippis, D. Ercolini, E. Parente, The microbiota of high-moisture mozzarella cheese produced with different acidification methods, International Journal of Food Microbiology 216 (2016) 9–17.
[12] P. Novielli, D. Romano, M. Magarelli, D. Diacono, A. Monaco, N. Amoroso, M. Vacca, M. De Angelis, R. Bellotti, S. Tangaro, Personalized identification of autism-related bacteria in the gut microbiome using explainable artificial intelligence, Iscience 27 (2024).
[13] M. Magarelli, P. Di Bitonto, F. De Filippis, P. Novielli, R. Magliulo, D. Diacono, R. Bellotti, S. Tangaro, Securing origin integrity through machine learning analysis of mozzarella di bufala pdo microbiome, in: 2024 IEEE International Workshop on Metrology for Industry 4.0 & IoT (MetroInd4. 0 & IoT), IEEE, 2024, pp. 38–42.
[14] F. De Filippis, V. Valentino, A. Alvarez-Ordóñez, P. D. Cotter, D. Ercolini, Environmental microbiome mapping as a strategy to improve quality and safety in the food industry, Current Opinion in Food Science 38 (2021) 168–176.
[15] T. Chen, C. Guestrin, Xgboost: A scalable tree boosting system, in: Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining, 2016, pp. 785–794.
[16] A. Chaudhary, S. Kolhe, R. Kamal, An improved random forest classifier for multi-class classification, Information Processing in Agriculture 3 (2016) 215–222.
[17] C. Schaffer, Selecting a classification method by cross-validation, Machine learning 13 (1993) 135–143.
[18] A. Holzinger, A. Saranti, C. Molnar, P. Biecek, W. Samek, Explainable ai methods-a brief overview, in: International workshop on extending explainable AI beyond deep models and classifiers, Springer, 2020, pp. 13–38.
[19] A. Lombardi, D. Diacono, N. Amoroso, P. Biecek, A. Monaco, L. Bellantuono, E. Pantaleo, G. Logroscino, R. De Blasi, S. Tangaro, et al., A robust framework to investigate the reliability and stability of explainable artificial intelligence markers of mild cognitive impairment and alzheimer's disease, Brain informatics 9 (2022) 17.
[20] M. Magarelli, P. Novielli, F. De Filippis, R. Magliulo, P. Di Bitonto, D. Diacono, R. Bellotti, S. Tangaro, Explainable artificial intelligence and microbiome data for food geographical origin: the mozzarella di bufala campana pdo case of study, Frontiers in Microbiology 15 (2024) 1393243.