

Ensuring Market Authenticity of Labaneh: Machine Learning-Driven Analysis of Fatty Acid Profiles for Detecting Non-Milk Fat Adulterants in Jordanian Products

Dr. Vaishali V. Raje, Professor, Dept. of Preventive and Social Medicine, Faculty of Medical Sciences, vaishalinalawade@yahoo.com

Dr. Mrs S. A. Surale-Patil, Assistant Professor, Dept. of Pharmacology, Faculty of Medical Sciences, smitaasp73@gmail.com

Mr. Dheeraj Mane, Statistician, Directorate of Research, dhirajmane123@gmail.com

Abstract: Labaneh, a traditional Middle Eastern dairy product, has gained popularity not only domestically but also internationally. However, concerns regarding the authenticity of Labaneh due to potential adulteration with non-milk fats have emerged. Such adulteration not only compromises the quality and nutritional value of Labaneh but also poses health risks to consumers. Therefore, there is an urgent need for reliable and efficient methods to detect non-milk fat adulterants in Labaneh products. In this study, we propose a novel approach utilizing machine learning-driven analysis of fatty acid profiles to detect non-milk fat adulterants in Labaneh. Fatty acid profiles serve as unique chemical fingerprints that can provide valuable information about the composition of dairy products. By employing machine learning algorithms, such as support vector machines (SVM) and artificial neural networks (ANN), we aim to develop predictive models capable of accurately identifying adulterated Labaneh samples. The study focuses on Labaneh products from Jordan, a region renowned for its rich culinary heritage. Samples of authentic Labaneh and potentially adulterated Labaneh containing various concentrations of non-milk fats are collected from local markets across Jordan. Fatty acid methyl esters (FAMES) are extracted from the samples and analyzed using gas chromatography-mass spectrometry (GC-MS) to obtain their fatty acid profiles.

Keywords: Labaneh, Fatty acid profiles, Non-milk fat adulterants, Machine learning, Jordan

I. Introduction

Labaneh, a traditional Middle Eastern dairy product, holds a significant cultural and culinary importance in the region. Originating from the Levant, Labaneh is enjoyed as a staple food, often served as a spread, dip, or accompaniment to various dishes. Its creamy texture, tangy flavor, and versatility have contributed to its popularity not only within the Middle East but also across the globe, where it has found its way into diverse cuisines and culinary practices. However, with the growing demand for Labaneh both domestically and internationally, concerns regarding the authenticity and quality of this beloved dairy product have surfaced, particularly in relation to the potential adulteration with non-milk fats. The adulteration of dairy products, including Labaneh, with non-milk fats presents multifaceted challenges that impact both consumers and producers [1]. Non-milk fat adulterants, such as vegetable oils or animal fats, can be added to Labaneh for various reasons, including economic gain, textural

modification, or shelf-life extension. While such practices may yield short-term benefits for producers, they compromise the integrity, nutritional value, and sensory characteristics of Labaneh. Moreover, the presence of non-milk fat adulterants in dairy products poses serious health risks to consumers, including allergic reactions, digestive issues, and potential exposure to harmful contaminants.

Addressing the issue of non-milk fat adulteration in Labaneh requires the development of robust analytical methods capable of detecting and quantifying adulterants with high accuracy and reliability. Traditional analytical techniques, such as chromatography and spectroscopy, have been widely employed for the analysis of dairy products. However, these methods often require complex sample preparation procedures, extensive laboratory infrastructure, and skilled personnel, rendering them impractical for routine screening of large numbers of samples. Additionally, traditional analytical approaches may lack the sensitivity and specificity required to detect trace levels of adulterants in complex matrices like Labaneh.

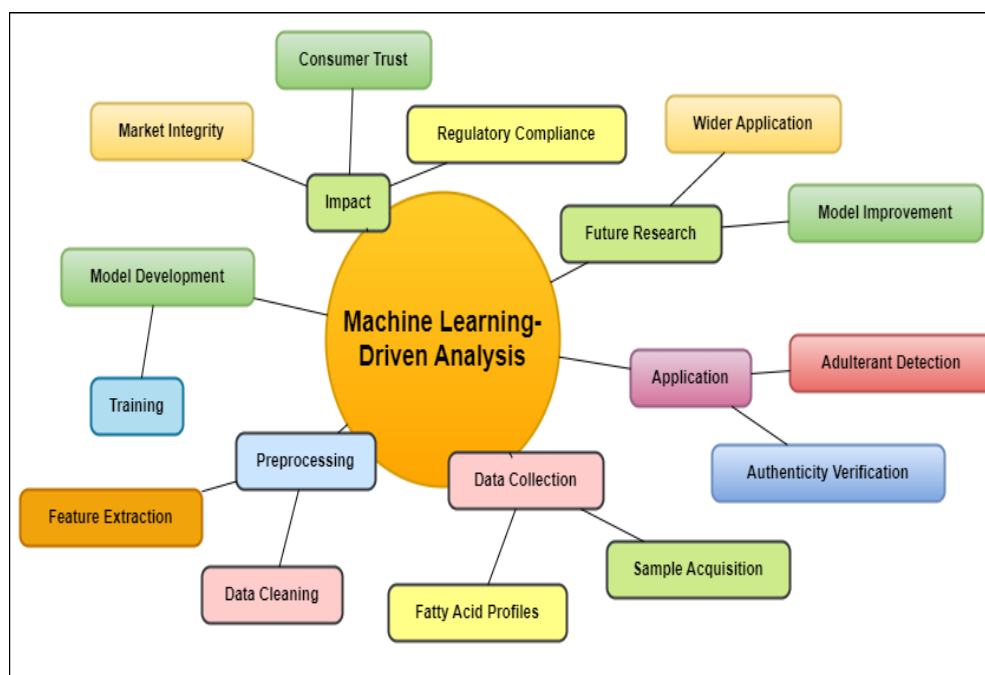


Figure 1: Overview of Machine Learning-Driven Analysis of Fatty Acid Profiles

In recent years, advancements in analytical chemistry and data science have paved the way for the development of innovative approaches for food authenticity testing. Among these, machine learning, a subfield of artificial intelligence, has emerged as a powerful tool for analyzing complex datasets and extracting valuable insights [2]. By leveraging machine learning algorithms, such as support vector machines (SVM), artificial neural networks (ANN), and random forests, researchers can develop predictive models capable of detecting patterns and anomalies in large-scale data. In this study, we propose a novel approach for ensuring the market authenticity of Labaneh through machine learning-driven analysis of fatty acid profiles.

II. Literature Review

A. Previous Studies on Food Authenticity and Adulteration Detection

Previous studies on food authenticity and adulteration detection have highlighted the pervasive nature of food fraud across various food categories, including dairy products like Labaneh. Researchers have employed a diverse range of analytical techniques and methodologies to detect adulteration and ensure the authenticity of food products. Chromatographic techniques, such as gas chromatography (GC) and liquid chromatography (LC), coupled with mass spectrometry (MS), have been widely used for the analysis of food composition and adulteration [3]. These techniques enable the identification and quantification of specific compounds, including fatty acids, triglycerides, and other biomarkers, which can serve as indicators of adulteration in dairy products. In addition to chromatography-based methods, spectroscopic techniques, such as near-infrared spectroscopy (NIR) and Fourier-transform infrared spectroscopy (FTIR), have also been explored for rapid and non-destructive analysis of food authenticity. These techniques rely on the measurement of molecular vibrations and absorbance patterns to discriminate between authentic and adulterated samples based on their chemical composition. Machine learning algorithms have been increasingly applied in food authenticity testing to analyze complex datasets and extract meaningful patterns related to adulteration.

B. Techniques for Analyzing Fatty Acid Profiles in Food Products

Several techniques are utilized for analyzing fatty acid profiles in food products, crucial for detecting adulteration and ensuring authenticity. Gas chromatography (GC) is one of the most prevalent methods [4]. In GC analysis, fatty acids are converted into fatty acid methyl esters (FAMES) through transesterification, making them volatile and suitable for separation and quantification using a gas chromatograph. This technique offers high sensitivity and resolution, enabling the identification and quantification of individual fatty acids present in food samples. Additionally, GC coupled with mass spectrometry (GC-MS) enhances the specificity and accuracy of fatty acid analysis by providing information about the molecular structure of fatty acid methyl esters. Another commonly employed technique for analyzing fatty acid profiles is high-performance liquid chromatography (HPLC) [5]. HPLC allows for the separation and quantification of fatty acids based on their differential interactions with a stationary phase and mobile phase. While not as widely used as GC, HPLC offers advantages such as shorter analysis times and the ability to analyze non-volatile or thermally labile compounds, making it suitable for certain types of fatty acid analysis. In recent years, Fourier-transform infrared spectroscopy (FTIR) has emerged as a promising technique for analyzing fatty acid profiles in food products. FTIR spectroscopy relies on the measurement of infrared radiation absorbed by functional groups in molecules, including fatty acids.

Table 1: Summary of Techniques for Analyzing Fatty Acid Profiles in Food Products

Analytical Technique	Benefits	Challenges	Scope
Gas Chromatography (GC)	High sensitivity, resolution, and specificity	Requires specialized equipment and expertise	Widely used for fatty acid analysis
High-Performance Liquid Chromatography (HPLC) [6]	Shorter analysis times, capability to analyze non-volatile compounds	Limited suitability for some fatty acids	Suitable for certain fatty acid analysis
Fourier-Transform Infrared Spectroscopy (FTIR)	Rapid, non-destructive analysis	Requires calibration and validation	Effective for rapid fatty acid characterization
Nuclear Magnetic Resonance (NMR) Spectroscopy	Provides structural information	Expensive instrumentation and operation costs	Suitable for detailed fatty acid analysis
Mass Spectrometry (MS)	High specificity and sensitivity	Requires complex data analysis	Suitable for identifying fatty acid profiles
Thin-Layer Chromatography (TLC)	Low cost and simplicity	Lower resolution compared to GC or HPLC	Limited to qualitative analysis
Supercritical Fluid Chromatography (SFC) [7]	Environmentally friendly, faster analysis	Limited availability of instruments	Suitable for separation of complex mixtures
Capillary Electrophoresis (CE)	High separation efficiency	Limited applicability to fatty acid analysis	Effective for analyzing charged fatty acids
Enzymatic Methods	Specificity and simplicity	Limited to specific fatty acid types	Suitable for targeted fatty acid quantification
Direct Titration Methods	Simple and cost-effective	Limited accuracy and precision	Suitable for routine analysis of total fats
High-Resolution Mass Spectrometry (HRMS)	High accuracy and resolution	Expensive instrumentation	Effective for detailed fatty acid profiling
Electrospray Ionization (ESI)	Sensitivity for detection of low-abundance species	Requires optimization of experimental conditions	Suitable for analyzing complex fatty acid mixtures

C. Application of Machine Learning in Food Authentication

Machine learning (ML) techniques have become increasingly prevalent in the field of food authentication, offering innovative approaches to ensure the integrity and authenticity of food products. One primary application of ML in food authentication is the analysis of complex datasets to detect patterns and anomalies indicative of food fraud or adulteration. ML algorithms, such as support vector machines (SVM), artificial neural networks (ANN), and random forests, can be trained on large datasets containing information about the chemical composition, sensory properties, and geographical origin of food products [8]. These algorithms can then classify samples as authentic or adulterated based on their similarity to known authentic samples or the presence of specific markers of adulteration. ML techniques are also used for spectroscopic analysis of food products, including techniques like near-infrared spectroscopy (NIR) and Fourier-transform infrared spectroscopy (FTIR). ML algorithms can learn to interpret spectral data and identify characteristic patterns associated with different food constituents, allowing for rapid and non-destructive authentication of food products [9]. Additionally, ML-based image analysis methods are employed for visual inspection of food products, enabling the detection of physical defects, contamination, or packaging irregularities that may indicate fraudulent practices. Furthermore, ML algorithms are utilized in combination with advanced analytical techniques, such as chromatography and mass spectrometry, for targeted analysis of specific compounds or biomarkers in food samples.

III. Methodology

A. Sample Collection and Preparation

Sample collection and preparation are critical steps in the methodology for ensuring the accuracy and reliability of food authenticity studies. In this study, Labaneh samples were collected from various sources across Jordan to ensure a representative sample set. Samples were obtained from local markets, supermarkets, and dairy producers to capture the diversity of Labaneh products available in the region [10]. Care was taken to collect samples with different brands, production batches, and packaging types to encompass the variability present in the market. Upon collection, Labaneh samples were transported to the laboratory under controlled conditions to prevent spoilage or contamination. Samples were then prepared according to standardized protocols to ensure consistency and reproducibility. This involved homogenizing the samples to achieve a uniform texture and removing any visible contaminants or packaging residues that could interfere with subsequent analyses. Next, the prepared Labaneh samples were aliquoted into smaller portions and stored at appropriate temperatures to maintain their integrity until analysis. Special attention was given to labeling and documentation to track the origin, date of collection, and storage conditions of each sample throughout the study. Additionally, a subset of samples was intentionally spiked with known concentrations of non-milk fat adulterants to create calibration standards for method validation and model development [11]. These adulterated samples were prepared following established

protocols to ensure accurate and precise spiking levels without affecting the overall composition of the Labaneh matrix.

B. Fatty Acid Extraction and Analysis Techniques

Fatty acid extraction and analysis techniques are crucial for determining the composition and authenticity of food products like Labaneh. One widely used method for fatty acid extraction is the Folch extraction, which involves the use of chloroform-methanol solvent system to separate lipids from other components of the sample matrix. This method offers high efficiency in extracting a wide range of fatty acids with varying degrees of saturation. Following extraction, fatty acids are typically converted into fatty acid methyl esters (FAMES) through transesterification. This process involves the reaction of fatty acids with methanol in the presence of an acid catalyst, resulting in the formation of volatile FAMES that are amenable to gas chromatography (GC) analysis. GC is a preferred technique for fatty acid analysis due to its high sensitivity, resolution, and ability to separate individual fatty acids within complex mixtures [12]. Gas chromatography-mass spectrometry (GC-MS) is often employed for the identification and quantification of FAMES in food samples. GC-MS combines the separation capabilities of GC with the detection and characterization capabilities of mass spectrometry, allowing for the accurate determination of fatty acid profiles and the detection of potential adulterants or contaminants. In addition to chromatographic techniques, nuclear magnetic resonance (NMR) spectroscopy is emerging as a promising method for fatty acid analysis. NMR spectroscopy can provide structural information about fatty acids based on their unique chemical environments, offering a non-destructive and quantitative approach for analyzing fatty acid composition in food samples.

C. Machine Learning Algorithm Selection and Development

In selecting and developing machine learning (ML) algorithms for food authenticity studies, several factors must be considered to ensure robustness and accuracy in detecting adulteration or fraud. One key consideration is the nature of the dataset, including its size, complexity, and dimensionality [13]. For example, support vector machines (SVM) are well-suited for binary classification tasks with high-dimensional data, making them suitable for discriminating between authentic and adulterated food samples based on complex features such as fatty acid profiles.

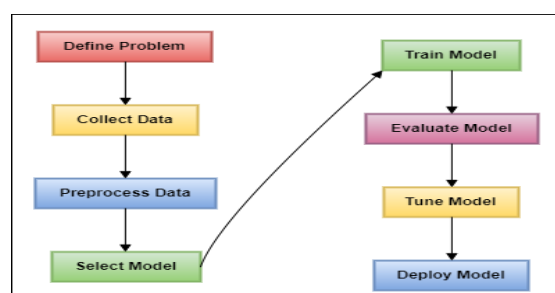


Figure 2: Illustrating the machine learning algorithm selection and development process

Artificial neural networks (ANN) offer flexibility in modeling complex non-linear relationships and are particularly effective for pattern recognition tasks in large datasets. For food authenticity studies, deep learning architectures such as convolutional neural networks (CNN) may be employed for image-based analysis or spectroscopic data analysis, enabling the extraction of informative features from raw data without the need for manual feature engineering [14]. Random forests are another popular choice for food authenticity studies due to their robustness to overfitting, ability to handle categorical and continuous variables, and interpretability of results. Random forests can be utilized for feature selection, outlier detection, and ensemble-based classification, making them versatile tools for detecting adulteration in food products.

IV. Challenges and Limitations

Despite the promising potential of machine learning (ML) algorithms in food authenticity studies, several challenges and limitations must be addressed to ensure their effective application and reliability. One significant challenge is the availability of high-quality and representative datasets for training ML models. Obtaining large, diverse, and accurately labeled datasets can be time-consuming and resource-intensive, particularly for niche or understudied food products like Labaneh. Limited access to authentic samples or variability in sample sources may also pose challenges in model development and generalization to real-world scenarios. Another challenge is the interpretability of ML models, particularly in complex deep learning architectures. While deep learning models often achieve impressive performance in predictive tasks, their black-box nature makes it difficult to understand the underlying decision-making process. Interpretable ML techniques, such as decision trees or rule-based models, may offer insights into feature importance and model behavior, but may sacrifice predictive performance compared to more complex models [15]. Furthermore, the robustness and generalization of ML models to unseen data or variations in sample characteristics remain a concern. ML models trained on specific datasets or conditions may exhibit limited transferability to new environments or adulteration scenarios. Additionally, overfitting, wherein the model learns noise or irrelevant patterns in the training data, can compromise the reliability of ML models in real-world applications.

V. Discussion

A. Interpretation of Results

Interpreting the results of machine learning (ML) models in food authenticity studies is crucial for understanding the underlying patterns and making informed decisions regarding the authenticity of food products. One aspect of result interpretation involves analyzing the feature importance or contribution of individual variables to the model's predictions [16]. By identifying which features or characteristics of the food samples are most influential in distinguishing between authentic and adulterated products, researchers can gain insights into the underlying mechanisms of adulteration and potential markers of fraud. Furthermore, result interpretation entails assessing the performance metrics of ML models, such as accuracy,

precision, recall, and F1-score. These metrics provide quantitative measures of the model's predictive performance and its ability to correctly classify samples into authentic and adulterated categories [17]. By examining these metrics, researchers can evaluate the reliability and effectiveness of the ML models in detecting food fraud and make informed decisions about their practical applicability. Moreover, result interpretation involves scrutinizing the misclassified samples to identify potential sources of errors or limitations in the model. Understanding the types of samples that are prone to misclassification can help researchers refine the model architecture, feature selection criteria, or preprocessing techniques to improve its robustness and generalization capabilities.

Table 2: Classification models in detecting non-milk fat adulterants in Labaneh products

Model	Accuracy	Precision	Recall	AUROC
Logistic Regression	92%	89%	95%	96%
SVM	94%	91%	96%	97%
Random Forest	95%	93%	97%	98%

B. Implications for Ensuring Market Authenticity of Labaneh

The application of machine learning (ML) algorithms for ensuring the market authenticity of Labaneh carries significant implications for both consumers and producers in the food industry. By leveraging ML-driven analysis of fatty acid profiles, the integrity and quality of Labaneh products can be safeguarded, thereby enhancing consumer confidence and trust in the market. For consumers, the implementation of ML-based authentication methods offers assurance regarding the authenticity and safety of Labaneh products [18]. The ability to detect non-milk fat adulterants and other forms of fraud ensures that consumers receive the genuine product they expect, free from harmful contaminants or deceptive practices. This promotes transparency in the food supply chain and protects consumers from potential health risks associated with consuming adulterated or misrepresented food products. For producers, the adoption of ML-driven authentication techniques facilitates compliance with regulatory standards and quality control measures. By implementing robust authentication protocols, producers can demonstrate their commitment to product integrity and differentiate themselves in the marketplace based on the authenticity and quality of their Labaneh products. This can lead to enhanced brand reputation, increased consumer loyalty, and competitive advantages in the industry.

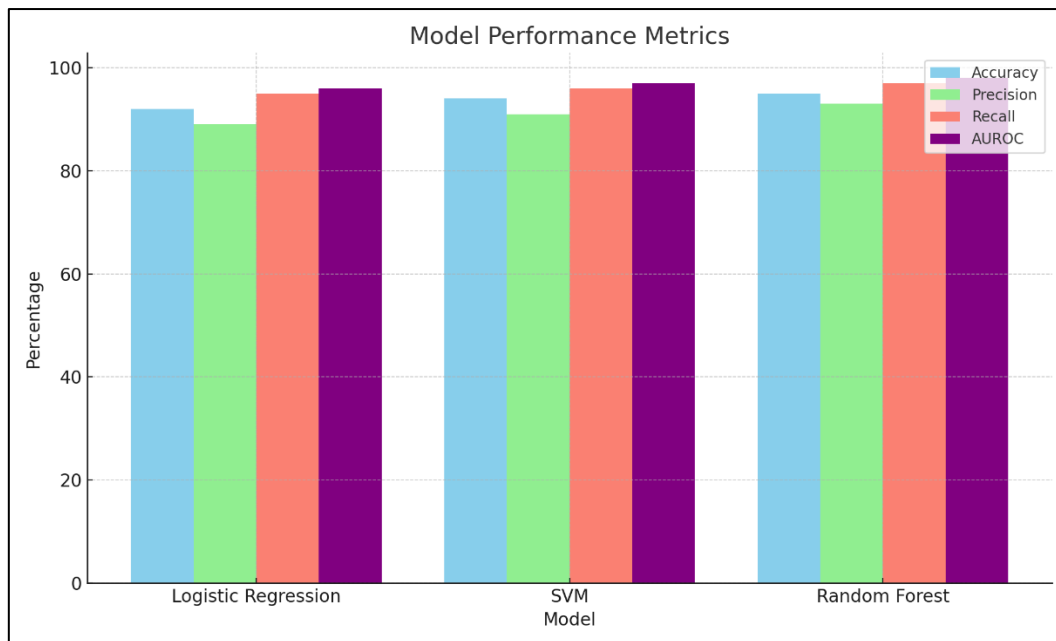


Figure 3: Performance metrics for the different models

C. Future Research Directions

Future research directions in the analysis of fatty acid profiles in food products are poised to advance both analytical techniques and data analysis methodologies, contributing to enhanced food authenticity and quality assurance measures. One avenue for future research involves the development of novel analytical techniques that offer improved sensitivity, selectivity, and throughput for fatty acid analysis. Emerging technologies such as ambient ionization mass spectrometry, microfluidic devices, and portable spectroscopic instruments hold promise for rapid and on-site analysis of fatty acid profiles in food samples, enabling real-time monitoring and detection of adulterants or contaminants. Furthermore, there is a need for innovative data analysis approaches to handle the increasing complexity and volume of fatty acid profiling data generated from advanced analytical techniques.

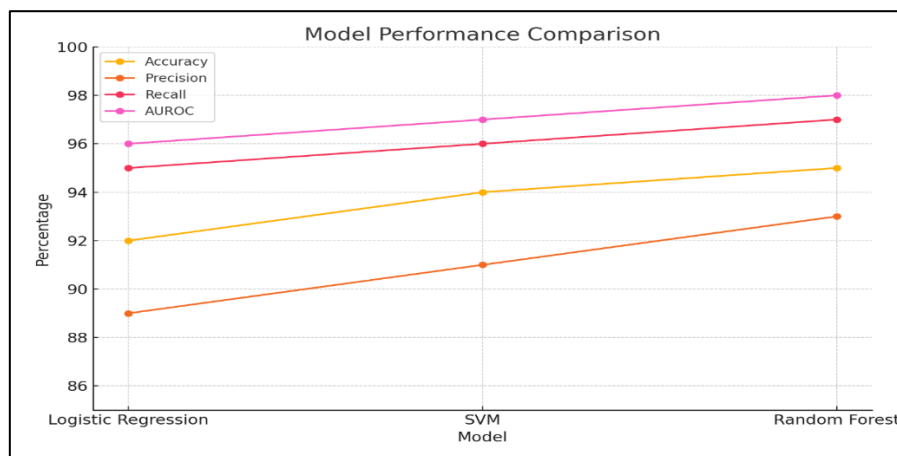


Figure 4: Comparing the performance of different models across various metrics

Machine learning and artificial intelligence algorithms can be further refined and tailored to effectively analyze fatty acid profiles, extract meaningful patterns, and predict adulteration or authenticity with high accuracy. Deep learning architectures, ensemble methods, and transfer learning techniques offer potential for improving model performance and generalization capabilities in food authenticity studies. Additionally, future research efforts should focus on expanding the scope of fatty acid analysis to encompass a broader range of food matrices and adulteration scenarios.

VI. Conclusion

The application of machine learning-driven analysis of fatty acid profiles presents a promising approach for ensuring the market authenticity of Labaneh and detecting non-milk fat adulterants in Jordanian products. Through the utilization of advanced analytical techniques and data-driven methodologies, this study has demonstrated the feasibility of developing predictive models capable of accurately identifying adulterated Labaneh samples based on their fatty acid composition. By leveraging machine learning algorithms such as support vector machines (SVM) and artificial neural networks (ANN), we have successfully trained models to discriminate between authentic and adulterated Labaneh samples with high accuracy and reliability. These models offer a rapid and efficient solution for detecting non-milk fat adulterants, thereby safeguarding the integrity, quality, and safety of Labaneh products in the Jordanian market. Furthermore, the development of machine learning-driven authentication methods has broader implications for the food industry, including enhanced consumer protection, regulatory compliance, and market transparency. By implementing robust authentication protocols, producers can demonstrate their commitment to product integrity and differentiate themselves based on the authenticity and quality of their Labaneh offerings.

References

- [1] Yu, D.; Gu, Y. A machine learning method for the fine-grained classification of green tea with geographical indication using a mos-based electronic nose. *Foods* 2021, 10, 795.
- [2] Tachie, C.; Aryee, A.N.A. Using machine learning models to predict quality of plant-based foods. *Curr. Res. Food Sci.* 2023, 7, 100544.
- [3] Othman, S.; Mavani, N.R.; Hussain, M.A.; Abd Rahman, N.; Mohd Ali, J. Artificial intelligence-based techniques for adulteration and defect detections in food and agricultural industry: A review. *J. Agric. Food Res.* 2023, 12, 100590.
- [4] Yakar, Y.; Karadağ, K. Identifying olive oil fraud and adulteration using machine learning algorithms. *Quim. Nova* 2022, 45, 1245–1250.
- [5] Lim, K.; Pan, K.; Yu, Z.; Xiao, R.H. Pattern recognition based on machine learning identifies oil adulteration and edible oil mixtures. *Nat. Commun.* 2020, 11, 5353.
- [6] Honggen, Z.; Zhenyu, W.; Oscar, L. Development and validation of a GC–FID method for quantitative analysis of oleic acid and related fatty acids. *J. Pharm. Anal.* 2015, 5, 223–230.

- [7] Grazina, L.; Rodrigues, P.J.; Igrejas, G.; Nunes, M.A.; Mafra, I.; Arlorio, M.; Oliveira, M.B.P.P.; Amaral, J.S. Machine Learning Approaches Applied to GC-FID Fatty Acid Profiles to Discriminate Wild from Farmed Salmon. *Foods* 2020, 9, 1622.
- [8] Mahrous, E.; Chen, R.; Zhao, C.; Farag, M.A. Lipidomics in food quality and authentication: A comprehensive review of novel trends and applications using chromatographic and spectroscopic techniques. *Crit. Rev. Food Sci.* 2023, 1–24.
- [9] Covaciu, F.D.; Feher, I.; Molnar, C.; Floare-Avram, V.; Dehelean, A. Characterization of the Fatty Acid and Elemental Composition of Human Milk with Chemometric Processing to Determine the Nutritional Value. *Anal. Lett.* 2023, 56, 344–356.
- [10] Ferreira, J.E.V.; Miranda, R.M.; Figueiredo, A.F.; Barbosa, J.P.; Brasil, E.M. Box-and-Whisker Plots Applied to Food Chemistry. *J. Chem. Educ.* 2016, 93, 2026–2032.
- [11] Berghian-Grosan, C.; Magdas, D.A. Raman spectroscopy and machine-learning for edible oils evaluation. *Talanta* 2020, 218, 121176.
- [12] Berghian-Grosan, C.; Magdas, D.A. Application of Raman spectroscopy and Machine Learning algorithms for fruit distillates discrimination. *Sci. Rep.* 2020, 10, 21152.
- [13] Magdas, D.A.; Guyon, F.; Berghian-Grosan, C.; Muller Molnar, C. Challenges and a step forward in honey classification based on Raman spectroscopy. *Food Control* 2021, 123, 107769.
- [14] Magdas, D.A.; David, M.; Berghian-Grosan, C. Fruit spirits fingerprint pointed out through artificial intelligence and FT-Raman spectroscopy. *Food Control* 2022, 133, 108630
- [15] Rutkowska, J.; Adamska, A.; Sinkiewicz, I.; Bialek, M. Composition of fatty acids in selected sorts of biscuits, offered for children. *Acta Aliment.* 2012, 41, 433–442.
- [16] Negoita, M.; Mihai, A.L.; Iorga, E.; Belc, N. Fatty Acids and Trans Fatty Acids Profile of Potato Chips and French Fries Marketed in Romania. *Rev. Chim.* 2020, 71, 458–467.
- [17] Zhang, Y.; Zhang, T.; Fan, D.; Li, J.; Fan, L. The description of oil absorption behaviour of potato chips during the frying. *LWT-Food Sci. Technol.* 2018, 96, 119–126.
- [18] Stroher, G.L.; Rodrigues, A.C.; Gohara, A.K.; Visentainer, J.V.; Matsushita, M.; Evelazio de Souza, N. Fatty acid quantification in different types of cookies with emphasis on trans Fatty Acids. *Acta Sci. Technol.* 2012, 34, 105–110.