

## Article

# Modelling Polyphenol Extraction through Ultrasound-Assisted Extraction by Machine Learning in *Olea europaea* Leaves

Raquel Rodríguez-Fernández, Ángela Fernández-Gómez, Juan C. Mejuto  and Gonzalo Astray \*

Universidade de Vigo, Departamento de Química Física, Facultade de Ciencias, 32004 Ourense, Spain; raquel.rodriguez.fernandez@uvigo.gal (R.R.-F.); angela.fernandez.gomez@alumnos.uvigo.gal (Á.F.-G.); xmejuto@uvigo.es (J.C.M.)

\* Correspondence: gastray@uvigo.es

**Abstract:** The study of the phenolic compounds present in olive leaves (*Olea europaea*) is of great interest due to their health benefits. In this research, different machine learning algorithms such as RF, SVM, and ANN, with temperature, time, and volume as input variables, were developed to model the extract yield and the total phenolic content (TPC) from experimental data reported in the literature. In terms of extract yield, the neural network-based ANN<sub>Z-L</sub> model presents the lowest root mean square error (RMSE) value in the validation phase (9.44 mg/g DL), which corresponds with a mean absolute percentage error (MAPE) of 3.7%. On the other hand, the best model to determine the TPC value was the neural network-based model ANN<sub>R</sub>, with an RMSE of 0.89 mg GAE/g DL in the validation phase (MAPE of 2.9%). Both models obtain, for the test phase, MAPE values of 4.9 and 3.5%, respectively. This affirms that ANN models would be good modelling tools to determine the extract yield and TPC value of the ultrasound-assisted extraction (UAE) process of olive leaves under different temperatures, times, and solvents.

**Keywords:** olive leaves; ultrasound-assisted extraction; extract yield; TPC; machine learning; random forest; support vector machine; artificial neural network



**Citation:** Rodríguez-Fernández, R.; Fernández-Gómez, Á.; Mejuto, J.C.; Astray, G. Modelling Polyphenol Extraction through Ultrasound-Assisted Extraction by Machine Learning in *Olea europaea* Leaves. *Foods* **2023**, *12*, 4483. <https://doi.org/10.3390/foods12244483>

Academic Editor: Hao Feng

Received: 7 November 2023

Revised: 5 December 2023

Accepted: 11 December 2023

Published: 14 December 2023



**Copyright:** © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Olive leaves (*Olea europaea*) are an interesting waste by-product of the olive oil industry [1]. Generally, they are acquired during olive harvesting or fabrication operations [2]. In the European Union (EU-28), oil mills generate 9.6 million tons/year of by-products that can be recovered, such as olive pomace, stone, and olive tree pruning biomass, which represents an additional 11.8 million tons [3]. Olive oil is the main lipid component of the Mediterranean diet [4]. As reported by Ben Hmida et al. (2022) [5], olive oil production and *Olea europaea* cultivation are significant and ancestral agricultural activities in Mediterranean countries. The EU produces around 67% of the world's total olive oil (mainly in the Mediterranean EU countries, representing about 4 million hectares [6]).

The olive plant has phenolic compounds that present beneficial health effects [7] and, as reported de Bock et al. (2013) [7], could possess antioxidant or anti-inflammatory properties, among others. Polyphenols are present principally in the tree's leaves and drupes [8]. The most ample phenolic compound in olive leaves is oleuropein [9]. As Martín García (2001) [10] reports in his Ph.D. thesis, the chemical composition of olive leaves, and therefore also their nutritional value, depends on a large number of factors such as climatic conditions, the variety of olive, the age of the plantation, or the harvest season. In fact, in the literature, studies can be found such as that by Cavalheiro et al. (2015) [11], in which the olive leaves of different varieties from Brazil are analysed according to moisture, ash, proteins, lipids, etc., or the research carried out by Ibrahim et al. (2016) [12], where different chemical compositions of whole and boiled leaves were studied.

On the other hand, Martín García (2001) [10] also reports that olive leaves generally have greater digestibility and a higher nutritional value than other by-products of oil extraction, but with a general observation that olive leaves have a low protein value.

Taking into account the trend to find abundant, renewable, and cheap sources of polyphenols [13], it can be understood that the characteristics that are present in olive leaves, together with the benefits that biophenols present, have given rise to a growing interest in their use in different industrial applications in the food supplement or pharmaceutical fields [2]. Nevertheless, the polyphenol profile in olive leaves is susceptible to several abiotic and biotic factors, such as the geographical zone, leaves, and so on [14]. On the other hand, there are certain important parameters that influence the extraction processes: the type and composition of the solvent, the extraction time, or the extraction temperature, among others [13]. Consequently, it is necessary to design and optimize the extraction methods for each polyphenol source [13].

Ultrasound-assisted extraction (UAE), as reported by Yerena-Prieto et al. (2022) [15], is a methodology based on cavitation, a phenomenon that causes the disruption of cell walls and having the effect that the release of target compounds [16]. In terms of the sound emitter devices (one of the parts of the core of ultrasonic equipment), two different types of devices can be used: an ultrasonic bath and a sonotrode (ultrasonic probe) [17]. As reported by Chahardoli et al. (2020) [18], UAE is a non-complex and reasonably priced method that can be applied at large scales for industrial and commercial purposes, can afford an improvement in mass transfer, provides an option for high-temperature procedures and a reduction in extraction time, and other economical or environmental benefits [19–22].

The high efficiency is caused by to the acoustic cavitation effect produced by the formation and implosion of microbubbles generated by ultrasonic waves [23].

The optimization and prediction of different properties in extraction methodologies can be approximated by traditional modelling methods such as response surface methodology (RSM) [13,24] or kinetic modelling [25]. Machine learning is an alternative to traditional methods; examples of this are the models based on random forest (RF), support vector machine (SVM), and artificial neural networks (ANNs).

According to Tian et al. (2017) [26], Breiman introduced RF for the first time in 2001 [27]. It is a classification and regression method that relies on statistical learning theory [26]. RF combines many prediction trees, where each tree is based on the values of an independently sampled random vector that presents the same distribution in all of the forest's trees [27]. In classification, the RF output corresponds to the class selected by a majority vote of all individual trees [26]; in regression, the predicted value is determined by averaging over all the trees [28]. RF can be utilized in different research fields, such as in agricultural science to detect nitrogen saturation [29], in environmental sciences to evaluate the exposure to particular matter in urban areas [30], and in food technology for quality control [31] and to detect food adulteration [32].

The next model, SVM, was first proposed, according to Waleed et al. (2020) [33], by Cortes and Vapnik in 1995 [34]. SVM is a supervised non-parametric statistical learning method; hence, it is not necessary to know the data distribution beforehand [35,36]. The operation is based on building a hyperplane to separate the data into different classes [35], while minimizing the classification error, through geometric margin maximization between classes [33]. SVM can be applied in different research areas, such as in chemistry to predict the toxicity of different compounds [37,38], in hydrology to determine water quality for drinking purposes [39], and in food technology to study plasticizers in extra virgin olive oil [40] or to classify Greek olive oils [41], among others.

Finally, the ANN technique is based on a computational system inspired by biological neural networks, which is composed of an input, an intermediate, and an output [42]. These networks consist of a set of artificial neurons that receive a series of signals (called inputs) that, once processed, if these signals exceed a threshold value, will be transmitted to another neuron [42]. According to Silva et al. (2015) [43], the ANN algorithm offers several advantages over conventional methods, such as its ability to fill in missing data,

its non-linearity, insensitivity to noise or high parallelism, among others [44]. ANNs have versatile applications across various fields of research, like in food technology for honey classification [45] or wine authentication [46], in chemistry to determine the electrical percolation of AOT-based microemulsions [47], and so on.

To summarize, in this study, different machine learning methods, (i) RF, (ii) SVM, and (iii) ANN, were applied to model the extract yield (mg/g DL)—DL corresponds to dried leaf—and the total phenolic content (TPC) (mg GAE/g DL) using experimental data obtained from the literature [25].

## 2. Materials and Methods

### 2.1. Experimental Data

The data used in this research were obtained from the experimental work carried out by Şahin (2019) [25]. According to Şahin (2019) [25], olive leaves were obtained from the Mediterranean region of Kaş (Turkey) during November in 2013, which is when the leaves have the highest content of phenolic content [48]. The leaves were dried under ambient conditions without exposure to light, and then stored in plastic bags in the absence of light until the time of grinding for extraction [25].

According to Şahin (2019) [25], UAE was performed using different solvent concentration ratios (aqueous solutions (*v/v*) at 50% of ethanol or methanol, and water, EtOH, or MeOH pure solutions) under a specific temperature (30 to 80 °C in steps of 10 °C for water; 30, 40, and 50 °C for EtOH, MeOH, and their aqueous solutions) and time range (20 to 60 min in steps of 10 min), with ninety being the total number of experimental cases.

The method developed by Malik and Bradford (2006) [49] was used by Şahin (2019) [25] to quantify the total polyphenolic content (TPC) present in the extract samples. A second-order kinetic model was used by Şahin (2019) [25] to carry out the kinetic description of the extraction.

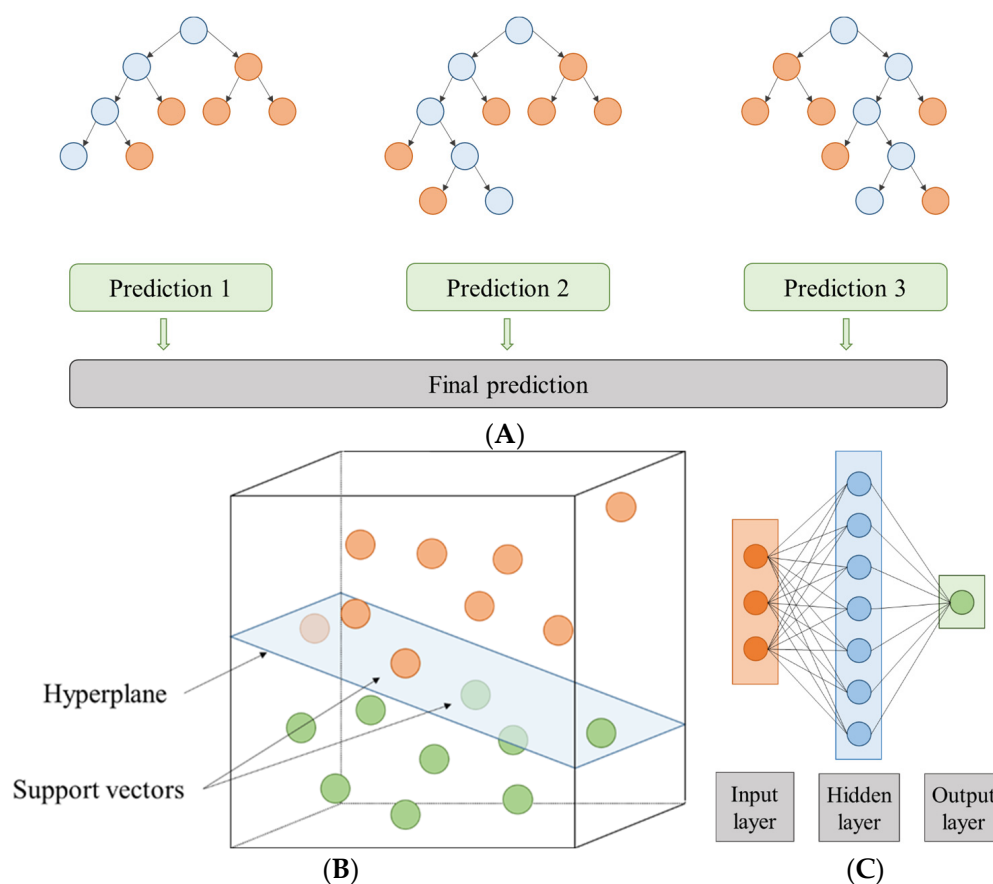
### 2.2. Machine Learning Approximations Developed

The database was divided into three groups: training (50%), validation (30%), and testing (20%). The training dataset, as the name suggests, was used to train the model, the validation group was used to find the best model, and the test group was retained to check if the model fit well with other external data. Temperature (°C), time (min), and volume (*v/v*) for each solvent type were used as input variables.

The first approach developed was random forest. The RF algorithm involves the random selection of samples from the original training dataset [50]. Its operation is based on a general classifier that is composed of several individual decision trees [51] (Figure 1A). Each of these individual classifiers is generated through a random vector that is taken independently from the input vector [51], and then their predictions are averaged [50]. This method has multiple benefits because the bootstrapping process decreases model variance without a bias increment [50,52]. It is also fast to analyse and is not vulnerable to overtraining and noise compared to other boosting-based techniques [50,52]. In this research, to find the best RF model, the following hyperparameters were tested: number of trees (in the range 1–100 using 99 steps), maximum depth (in the range 1–100 using 99 steps), attribute selection criteria (criterion least square), and pre-pruning (true or false). In addition, normalized RF models were also performed to fit the values to a specific range as the data had different units and scales. Two normalization methods were used: Z-transformation (indicated with subscript Z), which is a method of statistical normalization, and range transformation between −1 and 1 (indicated with subscript R). The normalization process was applied to the training group and, later, this normalization model was applied to the other data groups.

The SVM method is a classifier focused on splitting two linearly separable classes by searching for an optimal hyperplane [50,53] (Figure 1B). According to [50], to find the best hyperplane it is necessary to maximize the distance between the closest training cases (support vectors) and the separating hyperplane. To solve the SVM multiple classification

problem, as reported by Wang et al. (2020) [54], a combination of multiple binary classifiers is mainly used [55]. In this study, epsilon-SVR and nu-SVR were used to classify SVM types. In addition, C and gamma parameters were chosen to train the data. These parameters can be set according to the guide of Hsu et al. (2003) [56]. However, these working ranges were extended to improve model fits: the range of gamma values was approximately  $9.5 \times 10^{-7}$  to 256 with 28 steps and the range of C values was approximately  $9.8 \times 10^{-4}$  to 1,048,576 with 30 steps. The SVM library, proposed by [57], known as LibSVM, was used [58]. The two normalization methods, mentioned above, were also used: the Z-transformation and the range transformation between  $-1$  and  $1$ . In addition, the models were also applied in linear scale and logarithmic scale (indicated with subscript L) for gamma and C.



**Figure 1.** Scheme of (A) a random forest made up of three individual prediction trees, (B) a hyperplane in R3, and (C) an artificial neural network with a topology of 3-7-1.

ANN modelling is a simplified imitation of the human brain that it is composed of parallel processing units like the neurons present in living creatures [59]. In this study, the multilayer perceptron (MLP) was composed of input, intermediate, and output layers (each of these layers with a certain number of neurons) (Figure 1C). The input nodes depend on the experimental set, so in this case there were five (temperature, time, and volume % for each solvent type). In the intermediate layer, the neurons number was defined by  $2n + 1$ , with “n” being the number of input variables. Moreover, the neural models were developed with different hyperparameter configurations: training cycles (between 1 and 524,288 using 19 steps) and decay (true or false). In addition, the two normalization methods, Z-transformation and range transformation between  $-1$  and  $1$  scale, were carried out again on each of them, and linear scale and logarithmic scale were used for the training cycles.

General schemes for the RF, SVM, and ANN approaches that intended to model the extract yield and total phenolic content can be seen in Figure 1.

### 2.3. Best Model Selection

In this research, several statistical parameters were utilized to analyse the implemented machine learning models: (i) the root mean square error (RMSE), (ii) the mean absolute percentage error (MAPE in %), and (iii) the correlation coefficient ( $r$ ); these were determined for the training, validation, and testing groups. The equations can be consulted in Naeem et al. (2021) [60]. The model with the lowest RMSE value was considered the best model for each approach.

### 2.4. Equipment and Software

The obtained data were collected using Microsoft Excel 2013 (Microsoft, Redmond, WA, USA). The different models were created with RapidMiner Studio Educational 9.10.001 version (RapidMiner an ALTAIR Company, Troy, MI, USA). The computational equipment used was an Intel® Core™ i7-8700 CPU at 3.20 GHz with 64 GB of RAM and Windows 10 Pro. Figures were plotted with the SigmaPlot 13.0 (Systat Software Inc., San José, CA, USA).

## 3. Results and Discussion

In the following sections, corresponding to the Results and Discussion, the machine learning models selected based on the RMSE in the validation phase are detailed. Likewise, a discussion of the results is also carried out, comparing them with other research articles located in the literature.

### 3.1. Models to Yield Determination

Table 1 is shown below, presenting the models selected for each of the types of approximation carried out for the determination of yield and total polyphenol content.

**Table 1.** Adjustments for the selected algorithms to model the extract yield. RMSE corresponds to the root mean square error (mg/g), MAPE is the mean absolute percentage error (in %), and  $r$  corresponds to the correlation coefficient.

Model	Training			Validation			Test		
	RMSE	MAPE	$r$	RMSE	MAPE	$r$	RMSE	MAPE	$r$
RF	8.20	3.3	0.991	13.42	5.0	0.962	13.89	4.9	0.983
SVM <sub>L</sub>	8.36	2.9	0.986	9.87	3.9	0.979	5.95	2.6	0.997
ANN <sub>Z-L</sub>	2.71	0.9	0.999	9.44	3.7	0.980	12.82	4.9	0.985

For extract yield modelling, the first group of models developed was the random forest. Among this kind of approximation, the best RF model present an RMSE value of 13.42 mg/g DL in the validation phase that corresponds with a correlation coefficient of 0.962 (Table 1). The other two selected models developed (RF<sub>R</sub> and RF<sub>Z</sub>) showed slightly higher RMSE values (13.63 and 14.06 mg/g DL, respectively). The mean absolute percentage error is practically the same in all three models and varies around the range 5.0 to 5.5%. The selected RF model offers better adjustments for the training phase than for the validation phase. This can be seen in the different values obtained for the statistics under study, showing a clear decrease in the RMSE and the MAPE (8.20 mg/g DL and 3.3%) and an increase in the correlation coefficient value (0.991).

The second group of models (six SVM) were developed at the linear scale (SVM, SVM<sub>R</sub>, and SVM<sub>Z</sub>) and the logarithmic scale (SVM<sub>L</sub>, SVM<sub>R-L</sub>, and SVM<sub>Z-L</sub>). The RMSE results obtained from the SVM models at the logarithmic scale (9.87, 10.53, and 11.02 mg/g DL for SVM<sub>L</sub>, SVM<sub>R-L</sub>, and SVM<sub>Z-L</sub>, respectively) are significantly lower than the result obtained at the linear scale (14.16, 37.88, and 37.68 mg/g DL for SVM, SVM<sub>R</sub>, and SVM<sub>Z</sub>, respectively). Considering these results, it can be concluded that the best SVM model, considering the lowest RMSE error for the validation phase, is the SVM<sub>L</sub> model that presented an RMSE value of 9.87 mg/g DL (Table 1). These good adjustments in terms of RMSE are also reflected in the MAPE (3.9%) and in the correlation coefficient (0.979) obtained. These

good adjustments are also obtained in the training, where the model presents very similar values (RMSE of 8.36 mg/g DL and correlation coefficient of 0.986) to those reported for the validation.

Finally, the last kind of model was the artificial neural network. In this case, different ANNs were developed at the linear scale (ANN, ANN<sub>R</sub>, ANN<sub>Z</sub>) and the logarithmic scale (ANN<sub>L</sub>, ANN<sub>R-L</sub>, ANN<sub>Z-L</sub>). The root mean square error values obtained by these models in the validation phase varied between 14.16 mg/g DL (ANN<sub>L</sub>) and 9.44 mg/g DL (ANN<sub>Z-L</sub>). The two un-normalized models (ANN and ANN<sub>L</sub>) are the models with the highest RMSE in this phase (12.97 and 14.16 mg/g DL, respectively). The best model, as previously stated, is the ANN<sub>Z-L</sub> model, which presents a root mean square error of 9.44 mg/g DL for this phase, which corresponds to a MAPE of 3.7% (Table 1). The good behaviour in the validation phase is widely exceeded in the training phase, where the ANN<sub>Z-L</sub> model presents a relative RMSE of 2.71 mg/g DL and a mean absolute percentage error of 0.9%.

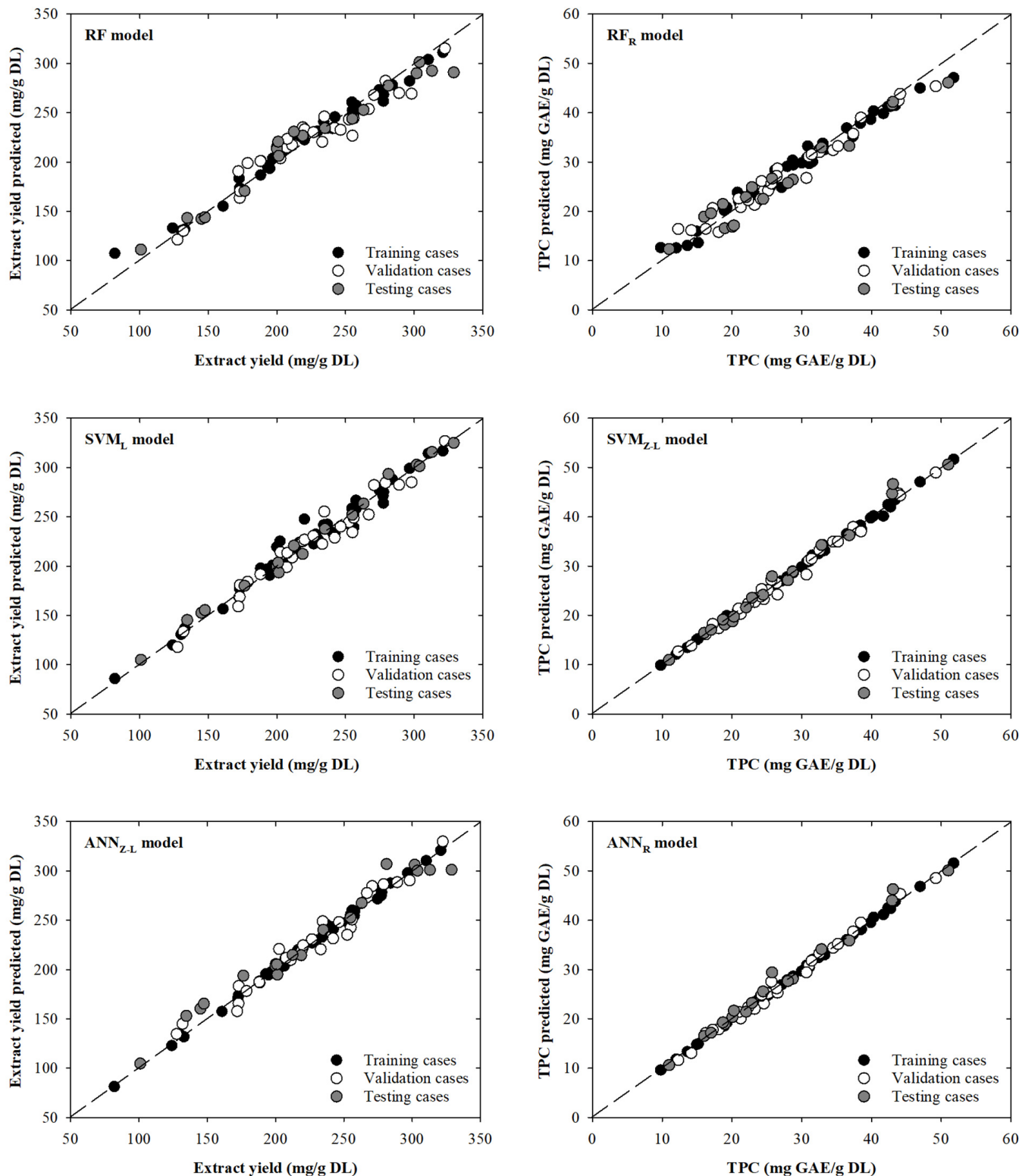
Among the three selected models chosen by the lowest RMSE in the validation phase (Table 1), the one with the worst adjustments in the validation phase is the RF model (13.42 mg/g DL), followed by the SVM<sub>L</sub> model (9.87 mg/g DL), which is very close to the best model, the ANN<sub>Z-L</sub> model (9.44 mg/g DL), developed with Z-transform normalization and logarithmic scale (ANN<sub>Z-L</sub>). These three models improve, for the training group, the adjustments obtained in the validation phase (especially the ANN model: 9.44 vs. 2.71 mg/g DL). It seems clear that the ANN<sub>Z-L</sub> model offers good results for both phases, training and validation, but it is necessary to verify its proper functioning in external data; for this, a previous group of data was reserved (test data). It can be seen that the ANN<sub>Z-L</sub> model offers good results based on RMSE (12.82 mg/g DL) and MAPE (4.9%).

Despite the fact that the ANN<sub>Z-L</sub> model offers a good fit for the test phase, it can be seen that one of the three previously selected models, the SVM<sub>L</sub>, offers better results for this phase. It can be seen that the root mean square error value is the lowest of the three selected models (5.95 mg/g DL); this also corresponds to the MAPE (2.6%).

Based on these results, it can be concluded that although the ANN<sub>Z-L</sub> model is, within the three selected models, the one that presents the lowest RMSE in the validation phase, and works correctly in the external validation phase, the SVM<sub>L</sub> model, whose statistics in the validation are very close to those obtained by the ANN<sub>Z-L</sub> model, is also a good performance modelling method due to the low error obtained in the test phase.

The RF model intended to model extract yield showed little dispersion between the experimental and modelled data. However, it can be seen that some cases move away from the line with slope one. For example, one training experimental case (81.9 mg/g DL) presents a significant deviation from the dashed line: an overestimation of 31.1% that corresponds with a modelled value of 107.4 mg/g DL. Additionally, another case (328.8 mg/g DL) in the testing phase presents a slightly deviation from the line with slope one, being modelled as 290.8 mg/g DL, which implies an underestimation of 11.6%. Based on these results, it can be observed that dispersion is mainly centred on extreme values of the extract yield values. This behaviour is not the case for the SVM<sub>L</sub> model, where dispersion is mainly concentrated in the middle–upper part of the graph. For instance, in one training case (220.0 mg/g DL), a value of 247.7 mg/g DL is modelled, resulting in an overestimation of 12.6%. The remaining cases in the training, validation, or testing phases are close to the line with slope one. This is clear in the test data that present better fits (RMSE = 5.95 mg/g DL) compared to the fits presented by the previous RF model (RMSE = 13.89 mg/g DL). On the other hand, as previously seen in the results shown in Table 1, the ANN<sub>Z-L</sub> model presents the lowest dispersion for the TVZ datasets. In fact, the model presents a MAPE value of 2.6% for all the data. Even so, in the upper part of the graph, there are two testing cases (281.2 and 328.8 mg/g DL) that are far from the dashed line with modelled values of 307.0 and 301.1 mg/g DL, resulting in an overestimation of 9.2% and an underestimation of 8.4%, respectively. Although these points are the most attractive to the human eye, there are a series of points towards the lower–middle zone of the graph that present a greater overestimation between 13.6% and 9.8%. All these

behaviours are reflected in Figure 2, where the scatter plots between the real value and those modelled by each of the selected models are shown. The points, for the extract yield models, are close, in general, to the black dashed line (line with slope one).



**Figure 2.** Scatter plots showing the real and modelled values of extract yield (left) and TPC (right) for the selected models. The black dashed line is the line with slope one.

### 3.2. Models for Total Phenolic Content Determination

The second group of models selected to model the TPC are shown in Table 2. The first approach was, again, a group of three random forests. In this sense, the best random forest model (RF<sub>R</sub>) presented an RMSE value of 1.93 mg GAE/g DL for the validation phase that corresponds with a correlation coefficient of 0.982. The other two models developed (RF and RF<sub>Z</sub>) showed similar RMSE values with minimum differences between them (2.00 and 1.94 mg GAE/g DL, respectively). Analysing the relative error, it remains unchanged across the three models, varying around 6.5% to 6.8%. The RF<sub>R</sub> model exhibits better adjustments during the training phase as opposed to the validation phase. These perceptible differences can be seen in the statistical parameters analysed: the RMSE and MAPE values slightly decreased (1.51 mg GAE/g DL and 4.6%), with an increase in the correlation coefficient value (0.990).

**Table 2.** Adjustments for the selected machine learning algorithms to model the total phenolic content. RMSE corresponds to the root mean square error (mg GAE/g DL), MAPE is the mean absolute percentage error (in %), and r corresponds to the correlation coefficient.

Model	Training			Validation			Test		
	RMSE	MAPE	r	RMSE	MAPE	r	RMSE	MAPE	r
RF <sub>R</sub>	1.51	4.6	0.990	1.93	6.5	0.982	2.47	9.4	0.978
SVM <sub>Z-L</sub>	0.41	0.8	0.999	0.93	2.7	0.995	1.23	3.1	0.995
ANN <sub>R</sub>	0.28	0.9	1.000	0.89	2.9	0.996	1.35	3.5	0.993

In addition to the RF models, a second group of six SVM models was also developed. The RMSE values obtained, 0.99, 0.95, and 0.93 mg GAE/g DL for SVM<sub>L</sub>, SVM<sub>R-L</sub>, and SVM<sub>Z-L</sub>, respectively, were significantly lower than for SVM, SVM<sub>R</sub>, and SVM<sub>Z</sub> (1.89, 7.19, and 7.35 mg GAE/g DL, respectively). Based on these results, it can be said that the SVM<sub>Z-L</sub> was the best model, obtaining an RMSE value of 0.93 mg GAE/g DL during the validation phase with a corresponding correlation coefficient of 0.995 and mean absolute percentage error of 2.7% (Table 2). Similar to the RF models, better fits were observed in the training phase than in the validation phase, with a decrease in RMSE and MAPE values (0.41 mg GAE/g DL and 0.8%, respectively) and an increase in the correlation coefficient value (0.999).

The final type of model included a set of six ANNs similar to the ANNs developed for extract yield. The un-normalized models in both scales (ANN and ANN<sub>L</sub>) showed slightly higher values of RMSE (1.52 and 1.70 mg GAE/g DL, respectively) compared to the normalized models (between 0.89 and 1.00 mg GAE/g DL, respectively). The best model was ANN<sub>R</sub>, which presented a root mean square error of 0.89 mg GAE/g DL in the validation phase that corresponds with a MAPE value of 2.9% and a correlation coefficient of 0.996 (Table 2). Consistent with previous findings, it can be observed that in the training phase the model shows better adjustment compared to the validation phase. In this case, the ANN<sub>R</sub> model presents a descent in both RMSE and MAPE (0.28 mg GAE/g DL and 0.9%, respectively).

Considering the three selected models based on the lowest RMSE during validation, the best model is ANN<sub>R</sub> (0.89 mg GAE/g DL), followed by SVM<sub>Z-L</sub> (0.93 mg GAE/g DL) and the RF<sub>R</sub> model (1.93 mg GAE/g DL). The correlation coefficients for all three models are high, greater than 0.980, and the relative errors are generally low (under 3.0%), except for the RF model, which exceeds 6.0%. All these models also show an improvement in the adjustment parameters for the training phase, especially the ANN<sub>R</sub> model, which improved from 0.89 mg GAE/g DL to 0.28 mg GAE/g DL. While the ANN<sub>R</sub> model appears to perform well in both phases, it is necessary to verify its performance using test data. In the testing phase, the ANN<sub>R</sub> model offers good results (1.35 mg GAE/g DL of RMSE), but it fails to achieve the best result. In this case, the SVM<sub>Z-L</sub> model performs better during testing with an RMSE of 1.23 mg GAE/g DL, a similarly high correlation coefficient (0.995



vs. 0.993), and a low MAPE value (3.1 vs. 3.5%). However, it can be considered that the differences between these two models, ANN<sub>R</sub> and SVM<sub>Z-L</sub>, are minimal.

Based on all the above, it can be said that the results show that ANN<sub>R</sub> has the lowest RMSE in the validation phase and works well in external validation. However, SVM<sub>Z-L</sub> has similar statistics and performs well in the test phase too, which makes both models good modelling methods.

Analysing the scatter plots presented in Figure 2, it can be seen that the RF<sub>R</sub> model designed to determine the TPC provides lower modelled values in the upper-right zone of Figure 2. Three specific cases of training, validation, and testing (51.8, 49.2, and 51.0 mg GAE/g DL), which decrease away from the dashed line (47.1, 45.4, and 46.1 mg GAE/g DL), present underestimations of 9.0, 7.8, and 9.6%. On the other hand, it can be seen that for the SVM<sub>Z-L</sub> and ANN<sub>R</sub> models, the dispersion around the trend line is smaller than for the RF<sub>R</sub> model. For the SVM<sub>Z-L</sub> and ANN<sub>R</sub> models, two testing cases can be highlighted (25.8 and 43.1 mg GAE/g DL), with overestimations of 8.4 and 8.3%, and 14.2 and 7.4%, respectively, for each model. By comparing the graphs and the statistical parameters obtained, it can be concluded that not only can the ANN<sub>R</sub> model be used, but also the SVM<sub>Z-L</sub> model could be used to model the variables of interest. This can be corroborated using the MAPE value for the total dataset (1.8%).

The results obtained in this research can be considered good for all the selected models. This statement is based on the good results obtained in all three phases where the selected models present MAPE values between 1.8 and 6.1%, which means that the percentage errors obtained by these models can be considered acceptable. The results obtained in this research can be compared with others reported in the literature.

In this sense, Alrugaibah et al. (2023) [61] carried out the extraction of different phenolic compounds, using deep eutectic solvents, by means of neural networks. The authors used four input variables and were able to obtain RMSE values between 0.04 and 0.16 mg/g [61]. Although the RMSE values are lower than those obtained in this research, it is necessary to indicate that the ranges studied for each of the phenolic compounds is lower [61] than the range studied in this research for the total extract yield. İlbay et al. (2014) [62] developed response surface methodology (RSM) and artificial neural network approaches to model and optimize the extraction process. In that work, the range of TPC (21.56 to 47.58 mg GAE/g DL) is similar to the range analysed in this research (9.79 to 51.78 mg GAE/g DL). In terms of RMSE, it can be observed that the ANN model presents a lower value (1.13 mg GAE/g DM) compared to the RSM (1.85 mg GAE/g DM) [62]. A similar work using response surface methodology was carried out by Goldsmith et al. (2014) [63] using temperature, time, and sample-to-solvent ratio as input variables, obtaining a model that can predict the TPC with an RMSE of 3.79 mg GAE/g for experimental values around 20 to 40 mg GAE/g. In the present study, the selected ANN model shows, for the total phases, an RMSE of 0.80 mg GAE/g DL, which indicates that the ANN model would work in a similar, or better, way than the models reported by İlbay et al. (2014) [62] and Goldsmith et al. (2014) [63].

More examples have also been found in the specialized literature; in this sense, Şahin et al. (2017) [64] developed RSM and ANN models to model and optimize TPC and oleuropein yields in olive leaf obtained by the SFMAE method (solvent-free microwave-assisted extraction). After an analysis of variance (ANOVA), the ANN model provides better results for TPC and oleuropein amount than the RSM model.

#### 4. Conclusions

In this research, three supervised machine learning approaches, random forest, support vector machine, and artificial neural network, were developed to model the extract yield and the total phenolic content extracted from *Olea europaea* leaves through ultrasound-assisted extraction. According to the results, it can be seen that the ANN models showed better performance in accuracy and generalization capability than the other models developed. Data normalization (Z-transformation and range transformation) was also identified

as an important factor to improve the effectiveness of the ANN. In this sense, the best selected neural network models present Z-transformation for the extract yield ANN model and range transformation for the total phenolic content ANN model. These two models provide good results in all phases, showing, for the querying phase, MAPEs of 4.9 and 3.5%, respectively. These findings suggest that ANN models using an appropriate normalization technique can be promising tools to model the extraction yield and TPC value. However, more research could be necessary to improve these results and their applicability in the pharmaceutical, food, or chemical industries.

Finally, it is necessary to emphasize that, with the increasing awareness of environmental issues and the growing need to reduce time, economic/material/personal costs, and the use of toxic products, the combination of these two methods, UAE and machine learning, will be an area of great interest among researchers that could lead to an improvement in the specialized industry.

**Author Contributions:** Conceptualization, G.A. and J.C.M.; methodology, R.R.-F., Á.F.-G. and G.A.; validation, R.R.-F. and G.A.; formal analysis, R.R.-F. and G.A.; investigation, R.R.-F. and G.A.; resources, J.C.M. and G.A.; writing—original draft preparation, R.R.-F., Á.F.-G. and G.A.; writing—review and editing, R.R.-F., Á.F.-G., J.C.M. and G.A.; visualization, R.R.-F., Á.F.-G. and G.A.; supervision, J.C.M. and G.A.; project administration, J.C.M. and G.A. All authors have read and agreed to the published version of the manuscript.

**Funding:** Raquel Rodríguez-Fernández thanks her employment contract “Actuación financiada por la Unión Europea, Next Generation UE 19.101.240A.482.96 Reserva de gestión directa. Servicio Público de Empleo Estatal. Inversión PAES III. Empleo Joven. Programa de Investigación. Mecanismo de Recuperación y Resiliencia mediante un contrato en la Universidad de Vigo”. Ángela Fernández-Gómez thanks the Ministerio de Educación y Formación Profesional for her grant 22CO1/008233 “Becas de colaboración de estudiantes en departamentos universitarios para el curso académico 2022-2023”.

**Data Availability Statement:** Data are contained within the article.

**Acknowledgments:** The authors thank RapidMiner Inc. for the educational and free license for RapidMiner Studio software (versions 9.10.001 and 9.10.011).

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Athanasiadis, V.; Grigorakis, S.; Lalas, S.; Makris, D.P. Highly Efficient Extraction of Antioxidant Polyphenols from *Olea europaea* Leaves Using an Eco-Friendly Glycerol/Glycine Deep Eutectic Solvent. *Waste Biomass Valorization* **2018**, *9*, 1985–1992. [CrossRef]
2. Şahin, S.; Bilgin, M. Olive Tree (*Olea europaea* L.) Leaf as a Waste by-Product of Table Olive and Olive Oil Industry: A Review. *J. Sci. Food Agric.* **2018**, *98*, 1271–1279. [CrossRef] [PubMed]
3. Berbel, J.; Posadillo, A. Review and Analysis of Alternatives for the Valorisation of Agro-Industrial Olive Oil By-Products. *Sustainability* **2018**, *10*, 237. [CrossRef]
4. Benavente-García, O.; Castillo, J.; Lorente, J.; Ortuño, A.; Del Rio, J.A. Antioxidant Activity of Phenolics Extracted from *Olea europaea* L. Leaves. *Food Chem.* **2000**, *68*, 457–462. [CrossRef]
5. Ben Hmida, R.; Gargouri, B.; Chtourou, F.; Sevim, D.; Bouaziz, M. Fatty Acid and Triacylglycerid as Markers of Virgin Olive Oil from Mediterranean Region: Traceability and Chemometric Authentication. *Eur. Food Res. Technol.* **2022**, *248*, 1749–1764. [CrossRef]
6. European Commission. Olive Oil in the EU. Available online: [https://agriculture.ec.europa.eu/farming/crop-productions-and-plant-based-products/olive-oil\\_es](https://agriculture.ec.europa.eu/farming/crop-productions-and-plant-based-products/olive-oil_es) (accessed on 30 October 2023).
7. de Bock, M.; Thorstensen, E.B.; Derraik, J.G.B.; Henderson, H.V.; Hofman, P.L.; Cutfield, W.S. Human Absorption and Metabolism of Oleuropein and Hydroxytyrosol Ingested as Olive (*Olea europaea* L.) Leaf Extract. *Mol. Nutr. Food Res.* **2013**, *57*, 2079–2085. [CrossRef]
8. Bucciantini, M.; Leri, M.; Nardiello, P.; Casamenti, F.; Stefani, M. Olive Polyphenols: Antioxidant and Anti-Inflammatory Properties. *Antioxidants* **2021**, *10*, 1044. [CrossRef]
9. El, S.N.; Karakaya, S. Olive Tree (*Olea europaea*) Leaves: Potential Beneficial Effects on Human Health. *Nutr. Rev.* **2009**, *67*, 632–638. [CrossRef]
10. Martín García, A.I. *Potencial de la Hoja de Olivo y del Orujo de dos Fases Como Alimentos Para Ovino y Caprino: Valoración Nutritiva Mediante Técnicas de Simulación*; Universidad de Granada: Granada, Spain, 2001.

11. Cavalheiro, C.V.; Picoloto, R.S.; Cichoski, A.J.; Wagner, R.; de Menezes, C.R.; Zepka, L.Q.; Da Croce, D.M.; Barin, J.S. Olive Leaves Offer More than Phenolic Compounds—Fatty Acids and Mineral Composition of Varieties from Southern Brazil. *Ind. Crops Prod.* **2015**, *71*, 122–127. [[CrossRef](#)]
12. Ibrahim, E.H.; Abdelgaleel, M.A.; Salama, A.A.; Metwalli, S.M. Chemical and Nutritional Evaluation of Olive Leaves and Selection the Optimum Conditions for Extraction Their Phenolic Compounds. *J. Sustain. Agric. Sci.* **2016**, *42*, 445–459. [[CrossRef](#)]
13. Şahin, S.; Şamlı, R. Optimization of Olive Leaf Extract Obtained by Ultrasound-Assisted Extraction with Response Surface Methodology. *Ultrason. Sonochem.* **2013**, *20*, 595–602. [[CrossRef](#)] [[PubMed](#)]
14. Talhaoui, N.; Taamalli, A.; Gómez-Caravaca, A.M.; Fernández-Gutiérrez, A.; Segura-Carretero, A. Phenolic Compounds in Olive Leaves: Analytical Determination, Biotic and Abiotic Influence, and Health Benefits. *Food Res. Int.* **2015**, *77*, 92–108. [[CrossRef](#)]
15. Yerena-Prieto, B.J.; Gonzalez-Gonzalez, M.; Vázquez-Espinosa, M.; González-de-Peredo, A.V.; García-Alvarado, M.Á.; Palma, M.; Rodríguez-Jimenes, G.D.; Barbero, G.F. Optimization of an Ultrasound-Assisted Extraction Method Applied to the Extraction of Flavonoids from Moringa Leaves (*Moringa oleifera* Lam.). *Agronomy* **2022**, *12*, 261. [[CrossRef](#)]
16. Mason, T.J.; Paniwnyk, L.; Lorimer, J.P. The Uses of Ultrasound in Food Technology. *Ultrason. Sonochem.* **1996**, *3*, S253–S260. [[CrossRef](#)]
17. Ranjha, M.M.A.N.; Irfan, S.; Lorenzo, J.M.; Shafique, B.; Kanwal, R.; Pateiro, M.; Arshad, R.N.; Wang, L.; Nayik, G.A.; Roobab, U.; et al. Sonication, a Potential Technique for Extraction of Phytoconstituents: A Systematic Review. *Processes* **2021**, *9*, 1406. [[CrossRef](#)]
18. Chahardoli, A.; Jalilian, F.; Memariani, Z.; Farzaei, M.H.; Shokohinia, Y. Chapter 26—Analysis of Organic Acids. In *Recent Advances in Natural Products Analysis*; Sanches Silva, A., Nabavi, S.F., Saeedi, M., Nabavi, S.M., Eds.; Elsevier: Amsterdam, The Netherlands, 2020; pp. 767–823. ISBN 978-0-12-816455-6.
19. Zhou, T.; Xu, D.-P.; Lin, S.-J.; Li, Y.; Zheng, J.; Zhou, Y.; Zhang, J.-J.; Li, H.-B. Ultrasound-Assisted Extraction and Identification of Natural Antioxidants from the Fruit of *Melastoma sanguineum* Sims. *Molecules* **2017**, *22*, 306. [[CrossRef](#)] [[PubMed](#)]
20. Rocha, J.C.G.; Procópio, F.R.; Mendonça, A.C.; Vieira, L.M.; Perrone, Í.T.; de Barros, F.A.R.; Stringheta, P.C. Optimization of Ultrasound-Assisted Extraction of Phenolic Compounds from Jussara (*Euterpe edulis* M.) and Blueberry (*Vaccinium myrtillus*) Fruits. *Food Sci. Technol.* **2018**, *38*, 45–53. [[CrossRef](#)]
21. Zu, G.; Zhang, R.; Yang, L.; Ma, C.; Zu, Y.; Wang, W.; Zhao, C. Ultrasound-Assisted Extraction of Carnosic Acid and Rosmarinic Acid Using Ionic Liquid Solution from *Rosmarinus officinalis*. *Int. J. Mol. Sci.* **2012**, *13*, 11027–11043. [[CrossRef](#)]
22. Soufi, O.; Medouni-Haroune, L.; Bachirbey, M.; Medouni-Adrar, S.; Idir, F.; Heddad, T.; Ouldsaadi, L.; Romero, C.; Madani, K.; Makhoulf-Boulekbache, L. Statistical Optimization of Ultrasound-Assisted Extraction of Polyphenols from Olive Pomace. *Sustain. Chem. Pharm.* **2023**, *36*, 101260. [[CrossRef](#)]
23. Arauzo, P.J.; Lucian, M.; Du, L.; Olszewski, M.P.; Fiori, L.; Kruse, A. Improving the Recovery of Phenolic Compounds from Spent Coffee Grounds by Using Hydrothermal Delignification Coupled with Ultrasound Assisted Extraction. *Biomass Bioenergy* **2020**, *139*, 105616. [[CrossRef](#)]
24. Giacometti, J.; Žauhar, G.; Žuvić, M. Optimization of Ultrasonic-Assisted Extraction of Major Phenolic Compounds from Olive Leaves (*Olea europaea* L.) Using Response Surface Methodology. *Foods* **2018**, *7*, 149. [[CrossRef](#)] [[PubMed](#)]
25. Şahin, S. Experimental and Modeling Study of Polyphenols in *Olea europaea* Leaves through Ultrasound-Assisted Extraction. *J. Turk. Chem. Soc. Sect. A Chem.* **2019**, *6*, 383–394. [[CrossRef](#)]
26. Tian, Y.; Yan, C.; Zhang, T.; Tang, H.; Li, H.; Yu, J.; Bernard, J.; Chen, L.; Martin, S.; Delepine-Gilon, N.; et al. Classification of Wines According to Their Production Regions with the Contained Trace Elements Using Laser-Induced Breakdown Spectroscopy. *Spectrochim. Acta Part B At. Spectrosc.* **2017**, *135*, 91–101. [[CrossRef](#)]
27. Breiman, L. Random Forests. *Mach. Learn.* **2001**, *45*, 5–32. [[CrossRef](#)]
28. Wang, L.; Zhou, X.; Zhu, X.; Dong, Z.; Guo, W. Estimation of Biomass in Wheat Using Random Forest Regression Algorithm and Remote Sensing Data. *Crop J.* **2016**, *4*, 212–219. [[CrossRef](#)]
29. Naicker, R.; Mutanga, O.; Peerbhay, K.; Agjee, N. The Detection of Nitrogen Saturation for Real-Time Fertilization Management within a Grassland Ecosystem. *Appl. Sci.* **2023**, *13*, 4252. [[CrossRef](#)]
30. Brokamp, C.; Jandarov, R.; Rao, M.B.; LeMasters, G.; Ryan, P. Exposure Assessment Models for Elemental Components of Particulate Matter in an Urban Environment: A Comparison of Regression and Random Forest Approaches. *Atmos. Environ.* **2017**, *151*, 1–11. [[CrossRef](#)]
31. Ribeiro, M.N.; Carvalho, I.A.; Fonseca, G.A.; Lago, R.C.; Rocha, L.C.R.; Ferreira, D.D.; Vilas Boas, E.V.B.; Pinheiro, A.C.M. Quality Control of Fresh Strawberries by a Random Forest Model. *J. Sci. Food Agric.* **2021**, *101*, 4514–4522. [[CrossRef](#)]
32. de Santana, F.B.; Borges Neto, W.; Poppi, R.J. Random Forest as One-Class Classifier and Infrared Spectroscopy for Food Adulteration Detection. *Food Chem.* **2019**, *293*, 323–332. [[CrossRef](#)]
33. Waleed, M.; Um, T.W.; Khan, D.A.; Khan, U. Automatic Detection System of Olive Trees Using Improved K-Means Algorithm. *Remote Sens.* **2020**, *12*, 760. [[CrossRef](#)]
34. Cortes, C.; Vapnik, V. Support-Vector Networks. *Mach. Learn.* **1995**, *20*, 273–297. [[CrossRef](#)]
35. Srestasathien, P.; Lawawirojwong, S.; Suwantong, R. Support Vector Regression for Rice Age Estimation Using Satellite Imagery. In Proceedings of the 2016 13th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON), Chiang Mai, Thailand, 28 June–1 July 2016; pp. 1–5.

36. Mountrakis, G.; Im, J.; Ogole, C. Support Vector Machines in Remote Sensing: A Review. *ISPRS J. Photogramm. Remote Sens.* **2011**, *66*, 247–259. [CrossRef]
37. Cao, D.-S.; Dong, J.; Wang, N.-N.; Wen, M.; Deng, B.-C.; Zeng, W.-B.; Xu, Q.-S.; Liang, Y.-Z.; Lu, A.-P.; Chen, A.F. In Silico Toxicity Prediction of Chemicals from EPA Toxicity Database by Kernel Fusion-Based Support Vector Machines. *Chemom. Intell. Lab. Syst.* **2015**, *146*, 494–502. [CrossRef]
38. Na, M.; Nam, H. Environmental Science Nano Predicting the Toxicity of Nano-Metal Oxide. *Environ. Sci. Nano* **2023**, *10*, 325–337. [CrossRef]
39. Panigrahi, N.; Patro, S.G.K.; Kumar, R.; Omar, M.; Ngan, T.T.; Giang, N.L.; Thu, B.T.; Thang, N.T. Groundwater Quality Analysis and Drinkability Prediction Using Artificial Intelligence. *Earth Sci. Inform.* **2023**, *16*, 1701–1725. [CrossRef]
40. Wu, X.; Ma, R.; Xu, B.; Wang, Z.; Du, Z.; Zhang, X.; Niu, Y.; Gao, S.; Liu, H.; Zhang, Y. Qualitative and Quantitative Studies of Plasticizers in Extra Virgin Olive Oil by Surface-Enhanced Raman Spectroscopy Combined with Chemometrics. *Vib. Spectrosc.* **2023**, *126*, 103527. [CrossRef]
41. Gyftokostas, N.; Nanou, E.; Stefan, D.; Kokkinos, V.; Bouras, C.; Couris, S. Classification of Greek Olive Oils from Different Regions by Machine Learning-Aided Laser-Induced Breakdown Spectroscopy and Absorption Spectroscopy. *Molecules* **2021**, *26*, 1241. [CrossRef]
42. Myronidis, D.; Ioannou, K. Forecasting the Urban Expansion Effects on the Design Storm Hydrograph and Sediment Yield Using Artificial Neural Networks. *Water* **2019**, *11*, 31. [CrossRef]
43. Silva, S.F.; Anjos, C.A.R.; Cavalcanti, R.N.; Celeghini, R.M. dos S. Evaluation of Extra Virgin Olive Oil Stability by Artificial Neural Network. *Food Chem.* **2015**, *179*, 35–43. [CrossRef]
44. Dębska, B.; Guzowska-Świder, B. Application of Artificial Neural Network in Food Classification. *Anal. Chim. Acta* **2011**, *705*, 283–291. [CrossRef]
45. Martínez-Castillo, C.; Astray, G.; Mejuto, J.C.; Simal-Gandara, J. Random Forest, Artificial Neural Network, and Support Vector Machine Models for Honey Classification. *EFood* **2020**, *1*, 69–76. [CrossRef]
46. Astray, G.; Martínez-Castillo, C.; Mejuto, J.-C.; Simal-Gandara, J. Metal and Metalloid Profile as a Fingerprint for Traceability of Wines under Any Galician Protected Designation of Origin. *J. Food Compos. Anal.* **2021**, *102*, 104043. [CrossRef]
47. Montoya, I.A.; Moldes, Ó.; Cid Samamed, A.; Astray, G.; Galvez, J.F.; Mejuto, J.C. Influence Prediction of Alkylamines Upon Electrical Percolation of AOT-Based Microemulsions Using Artificial Neural Networks. *Tenside Surfactants Deterg.* **2015**, *52*, 473–476. [CrossRef]
48. Şahin, S.; Sa, N.; Perez, J.; Brockington, J. Seasonal Changes of Individual Phenolic Compounds in Leaves of Twenty Olive Cultivars Grown in Texas. *J. Agric. Sci. Technol.* **2012**, *2*, 242–247.
49. Malik, N.S.A.; Bradford, J.M. Changes in Oleuropein Levels during Differentiation and Development of Floral Buds in ‘Arbequina’ Olives. *Sci. Hortic.* **2006**, *110*, 274–278. [CrossRef]
50. Rodríguez-Garlito, E.C.; Paz-Gallardo, A. Efficiently Mapping Large Areas of Olive Trees Using Drones in Extremadura, Spain. *IEEE J. Miniaturizat. Air Space Syst.* **2021**, *2*, 148–156. [CrossRef]
51. Moriondo, M.; Stefanini, F.M.; Bindi, M. Reproduction of Olive Tree Habitat Suitability for Global Change Impact Assessment. *Ecol. Modell.* **2008**, *218*, 95–109. [CrossRef]
52. Ghamisi, P.; Plaza, J.; Chen, Y.; Li, J.; Plaza, A.J. Advanced Spectral Classifiers for Hyperspectral Images: A Review. *IEEE Geosci. Remote Sens. Mag.* **2017**, *5*, 8–32. [CrossRef]
53. Aguilera Puerto, D.; Cáceres Moreno, Ó.; Martínez Gila, D.M.; Gómez Ortega, J.; Gámez García, J. Online System for the Identification and Classification of Olive Fruits for the Olive Oil Production Process. *J. Food Meas. Charact.* **2019**, *13*, 716–727. [CrossRef]
54. Wang, X.; Wang, G.; Hou, X.; Nie, S. A Rapid Screening Approach for Authentication of Olive Oil and Classification of Binary Blends of Olive Oils Using Low-Field Nuclear Magnetic Resonance Spectra and Support Vector Machine. *Food Anal. Methods* **2020**, *13*, 1894–1905. [CrossRef]
55. Hsu, C.W.; Lin, C.J. A Comparison of Methods for Multiclass Support Vector Machines. *IEEE Trans. Neural Netw.* **2002**, *13*, 415–425. [CrossRef] [PubMed]
56. Hsu, C.; Chang, C.; Lin, C.-J. A Practical Guide to Support Vector Classification. Available online: <https://www.csie.ntu.edu.tw/~cjlin/> (accessed on 30 March 2023).
57. Chang, C.-C.; Lin, C.-J. LIBSVM: A Library for Support Vector Machines. *ACM Trans. Intell. Syst. Technol.* **2011**, *2*, 1–27. [CrossRef]
58. RapidMiner Support Vector Machine (LibSVM). Available online: [https://docs.rapidminer.com/latest/studio/operators/modeling/predictive/support\\_vector\\_machines/support\\_vector\\_machine\\_libsvm.html](https://docs.rapidminer.com/latest/studio/operators/modeling/predictive/support_vector_machines/support_vector_machine_libsvm.html) (accessed on 30 March 2023).
59. Gonzalez-Fernandez, I.; Iglesias-Otero, M.A.; Esteki, M.; Moldes, O.A.; Mejuto, J.C.; Simal-Gandara, J. A Critical Review on the Use of Artificial Neural Networks in Olive Oil Production, Characterization and Authentication. *Crit. Rev. Food Sci. Nutr.* **2019**, *59*, 1913–1926. [CrossRef] [PubMed]
60. Naem, M.; Yu, J.; Aamir, M.; Khan, S.A.; Adeleye, O.; Khan, Z. Comparative Analysis of Machine Learning Approaches to Analyze and Predict the COVID-19 Outbreak. *PeerJ. Comput. Sci.* **2021**, *7*, e746. [CrossRef] [PubMed]
61. Alrugaiabah, M.; Yagiz, Y.; Gu, L. Novel Natural Deep Eutectic Solvents as Efficient Green Reagents to Extract Phenolic Compounds from Olive Leaves and Predictive Modelling by Artificial Neural Networking. *Food Bioprod. Process.* **2023**, *138*, 198–208. [CrossRef]

62. İlbay, Z.; Şahin, S.; Büyükkabasakal, K. A Novel Approach for Olive Leaf Extraction through Ultrasound Technology: Response Surface Methodology versus Artificial Neural Networks. *Korean J. Chem. Eng.* **2014**, *31*, 1661–1667. [[CrossRef](#)]
63. Goldsmith, C.D.; Vuong, Q.V.; Stathopoulos, C.E.; Roach, P.D.; Scarlett, C.J. Optimization of the Aqueous Extraction of Phenolic Compounds from Olive Leaves. *Antioxidants* **2014**, *3*, 700–712. [[CrossRef](#)]
64. Şahin, S.; Samli, R.; Tan, A.S.B.; Barba, F.J.; Chemat, F.; Cravotto, G.; Lorenzo, J.M. Solvent-Free Microwave-Assisted Extraction of Polyphenols from Olive Tree Leaves: Antioxidant and Antimicrobial Properties. *Molecules* **2017**, *22*, 1056. [[CrossRef](#)]

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.