Talanta 161 (2016) 304–308

Contents lists available at ScienceDirect

Talanta

journal homepage: www.elsevier.com/locate/talanta

Neural networks applied to characterize blends containing refined and extra virgin olive oils

Regina Aroca-Santos, John C. Cancilla, Enrique S. Pariente, José S. Torrecilla *

Departamento de Ingeniería Química, Facultad de Ciencias Químicas, Universidad Complutense de Madrid, Madrid 28040, Spain

ARTICLE INFO

Article history:
Received 7 May 2016
Received in revised form
7 August 2016
Accepted 10 August 2016
Available online 10 August 2016

Keywords:
Extra virgin olive oil varietals
Visible spectroscopy
Artificial neural networks
Binary blends

ABSTRACT

The identification and quantification of binary blends of refined olive oil with four different extra virgin olive oil (EVOO) varietals (Picual, Cornicabra, Hojiblanca and Arbequina) was carried out with a simple method based on combining visible spectroscopy and non-linear artificial neural networks (ANNs).

The data obtained from the spectroscopic analysis was treated and prepared to be used as independent variables for a multilayer perceptron (MLP) model. The model was able to perfectly classify the EVOO varietal (100% identification rate), whereas the error for the quantification of EVOO in the mixtures containing between 0% and 20% of refined olive oil, in terms of the mean prediction error (MPE), was 2.14%. These results turn visible spectroscopy and MLP models into a trustworthy, user-friendly, low-cost technique which can be implemented on-line to characterize olive oil mixtures containing refined olive oil and EVOOs.

© 2016 Elsevier B.V. All rights reserved.

1. Introduction

Over the years, research grows exponentially and becomes more complex and intricate as time passes. This fact has originated the requirement of new and efficient experimental and data treating techniques. Specifically, in relation to data analysis, linear and non-linear chemometric approaches have been present for a long time and they will continue being needed in the scientific community all around the globe, as they are in many cases the final step of a research study. This work will focus on artificial neural networks (ANNs) as the main mathematical modeling tool. In the recent past, it has been employed in significant fields such as food technology [1,2], nanotechnology [3], and chemistry among others [4], which shows the wide applicability of such tools.

ANNs differ from traditional linear methods in their ability to adequately find non-linear relationships between the analyzed variables within a database. This main characteristic is also an advantage as it qualifies these tools to work with different samples simultaneously, enabling, for example, the analysis of data of different olive oil blends with one single ANN-based model (as required for the present task).

Oil blending, fraudulent or accidental, is nowadays an issue which needs to be taken care of due to both financial and health reasons, to avoid tragedies such as the toxic oil syndrome (TOS) back in 1981, which led to many human deaths [5]. Also, protecting the reputation of high quality extra virgin olive oils (EVOOs) is of great importance, especially for main producing countries such as Spain, Italy, or Greece.

EVOO is well known for being a vital ingredient of the Mediterranean diet [6]. This prestigious functional food offers beneficial effects on human health such as aiding in the prevention of cardiovascular diseases [7,8], breast cancer [9], and oxidative stress related sicknesses among others [10,11]. In light of these wonderful properties, the quality of EVOO is a protected good for producing countries and, therefore, there are many analytical techniques for olive oil characterization, authentication, and adulterant detection, such as: solid phase microextraction coupled with gas chromatography for the analysis of volatile compounds [12], FTIR spectroscopy [13], near infrared spectroscopy [14], nuclear magnetic resonance [15], high performance liquid chromatography [16], synchronous fluorescence spectroscopy [17], and visible spectroscopy [1]. Out of all of these analytical techniques, visible spectroscopy was chosen to carry out the measurements of 80 binary blends of refined olive oil with four EVOOs of different varietals (Picual, Cornicabra, Hojiblanca and Arbequina), because it is a fast, straightforward, and reliable technique [18,19].

To sum up, the goal of this work is to combine visible spectroscopy with ANNs to identify the EVOO contained in the blend and the volume percentage of refined olive oil present in each sample.

2. Materials and methods

In this section, the blends used and the analytical and data

* Corresponding author.
E-mail address: jstorre@ucm.es (J.S. Torrecilla).

http://dx.doi.org/10.1016/j.talanta.2016.08.033
0039-9140/© 2016 Elsevier B.V. All rights reserved.

treatment methods employed will be explained.

2.1. Olive oil blends

Binary blends of refined olive oil with Arbequina, Cornicabra, Hojiblanca, and Picual varietals were prepared within the whole volume percentage range (0–100%) in intervals of 5%, leading to a total of 80 oil mixtures ready to be measured without any other pretreatment, meaning that no solvents were used. This last trait qualifies the method for its on-line use.

The oils used (shown in Table 1) were within their best before dates and they were properly stored away from direct light and at room temperature.

2.2. Vis-spectroscopy

To carry out the measurements of the blends, a Varian Cary 1E UV–vis spectrophotometer operating between 350 and 750 nm was employed. The blends were analyzed in a 1 cm path length quartz cell. Each blend was measured three times and the average of the measurements was used as the final spectrum of each sample.

The area under the curves (AUCs) of the representative peaks of the spectra obtained were then used as independent variables during the modeling phase, as it will be explained further on.

2.3. Artificial neural networks

The mathematical tool proposed for the determination of EVOOs and the quantification of refined olive oil in the binary blends measured, was an ANN. These are non-linear models capable of satisfactorily finding non-linear relationships between the different independent and dependent variables present in the analyzed database (in this case absorption and the concentrations of the components in the blends used, respectively), making them suitable for working with different samples (olive oil blends in this case) at the same time. The estimating mechanism of ANNs is based on interpolating, so it is important to highlight that they strictly depend on the range of the existing data and they do not provide reliable or accurate extrapolations [20].

The chosen ANN for this particular case was a multilayer perceptron (MLP)-based model, due to its successful implementation to carry out numerous applications in many fields such as chemistry [4], nanochemistry [3], biochemistry [21], and food technology [19]. A MLP is a supervised model, which means that its learning and optimizing processes require labeled data (input and its respective target data) [22].

MLPs are structured in three different layers as shown in Fig. 1: the input layer containing the nodes in charge of introducing all the data into the MLP, and the hidden and output layers, comprised of neurons, which are the calculation centers responsible of converting the input signals into processed output responses [23].

The number of units in the input and the output layer depend on the case studied, as they are the independent and dependent variables, respectively. On the other hand, the number of neurons in the hidden layer is an important parameter which must be thoughtfully considered when optimizing a MLP, because an excessive number of these units would create an over-fit model, whereas an insufficient amount would lead to a model with a weak learning ability [20,24]. All of the elements in one layer are connected to every element in the next layer by weighted parameters called “weights”, which are of vital importance in the optimization of the MLP, as its precision and performance mainly depend on their correct optimization [22,25,26].

Table 1
Oils used for the measured blends.

Type	Supplier
Arbequina varietal EVOO	Borges Pont S.A.U. Oils
Cornicabra varietal EVOO	Toledo, S.A., Oils
Hojiblanca varietal EVOO	Borges Pont S.A.U. Oils
Picual varietal EVOO	Borges Pont S.A.U. Oils
Refined olive oil (ROO)	KOIPE, SOS Cuétara SA

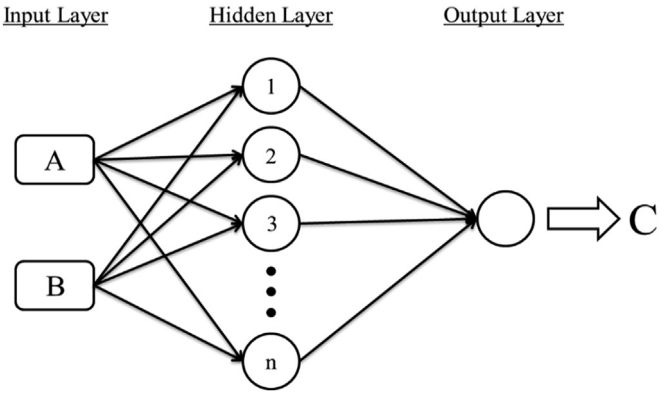

Fig. 1. Scheme of a multilayer perceptron, where A and B are independent input variables and C is the dependent output variable.

2.3.1. Optimization of the MLP

The optimization of ANN-based models is necessary to accomplish a precise, trustworthy, and functional estimating tool [18,27]. This optimization involves two different steps: a training and a verification process. To carry out these two steps properly, the original database must be randomly divided in two groups, the training (80% of the data) and the verification (20% of the data) datasets. This division must be done considering that the MLP requires sufficient information to train the network properly and, at the same time, it needs a representative quantity of samples to verify its generalization capability [22,23]. Throughout the training phase, the input information confined in the training dataset is presented to the model and an estimated value is obtained. This answer is then compared with the target value, and the error is back-propagated through the network, permitting the weight modification to lower this error [27,28]. If this process does not stop, the training error would asymptotically tend to zero, attaining a faultless fit for the training data. This would originate an over-fit system, incapable of delivering accurate estimations for external data. In order to avoid this, the verification step takes place. The verification data is used to analyze the generalizing capability of the model, and here the values of the weights are not modified; they stay fixed, which eventually leads to an error increase if the training cycles continue [20]. When the verification error rises for a series of six consecutive cycles, the training process is stopped, and the weights are considered optimized. Thus, the objective of both processes (training and verification steps) is to raise the accuracy of the MLP and to decrease the prediction error (in terms of mean prediction error, MPE), Eq. (1) for data that is not contained in the training dataset [18,28].

$$MPE = \frac{1}{N}\sum_{k=1}^{N}\frac{|r_k - y_k|}{r_k} \times 100 \quad (1)$$

Where $\boldsymbol{N}$ represents the number of data points evaluated, $\boldsymbol{r_k}$ is the real value and $\boldsymbol{y_k}$ is the value estimated by the MLP-based model.

Following this key optimization process, there are other intrinsic parameters of the ANN-based model that need additional consideration and optimization before the model is completely optimized and ready to perform. These parameters are the training

Download English Version:

https://daneshyari.com/en/article/1241718

Download Persian Version:

https://daneshyari.com/article/1241718

Daneshyari.com