Assessing the Performance of a Machine Learning System to Predict Geometrical Properties of Ahmad Aghaei Pistachio Kernels

Fatemeh Koushki(MSc)[1], Hamid Tavakolipour(PhD)[2], Mohsen Mokhtarian(PhD)[1*]

[1] *Department of Food Science and Technology, Roudehen Branch, Islamic Azad University, Roudehen, Iran*
[2] *Department of Food Science and Technology, Sabzevar Branch, Islamic Azad University, Sabzevar, Iran.*

Information

Article Type:
Original Article

Article History:

Received: 10.01.2022
Accepted: 25.02.2022

Doi: 10.22123/PHJ.2022.324121.1120

Keywords:
Pistachio (Pistacia Vera L.)
Artificial Neural Network
Machine Learning System
Modeling and Predicting Engineering Properties.

Corresponding Author:
Mohsen Mokhtarian

Email: mokhtarian.mo@riau.ac.ir

Tel: +98- 9352601788

Abstract

Background: The use of machine learning techniques such as artificial neural networks (ANN) improves the performance and speed of prediction processes as well as their reliability in the design of agricultural processing machines. Machine learning as a subset of artificial intelligence makes it possible to develop a unique way to create a predictive model system in the form of a known dataset by developing machine learning models (MLM).

Materials and Methods: In this study, first the geometric properties of pistachio kernels including the major diameter (L), intermediate diameter (T), minor diameter (W), geometric mean diameter (D_g), and surface area (S) were calculated at four moisture levels of 4.33, 22.64, 29.11, and 41.35% (w.b). Then, the data obtained in this step were used as the input values (L, W & T) and the output value (S) into the machine learning system. Multi-layer perceptron (MLP) and radial basis functions (RBF) were used as two machine learning models to predict the surface area of pistachio kernel during rehydration.

Results: The data analysis revealed that the neural network model of RBF with 42 neurons in the hidden layer (N_{1st}=42) had the lowest mean relative error (MRE=0.01414), and the highest coefficient of determination (R^2=0.954) and chosen as the best model for predicting the surface area of pistachio kernel.

Conclusion: Following the findings of this study, it can be concluded that the MLM as one of new forecasting techniques can be used to estimate the engineering properties of agricultural products.

►Please cite this article as follows:
Koushki F, Tavakolipour H, Mokhtarian M. Assessing the Performance of a Machine Learning System to Predict Geometrical Properties of Ahmad Aghaei Pistachio Kernels. Pistachio and Health Journal. 2022; 5(1): 22-29 .

1. Introduction

Pistachio (*Pistacia vera* L.) is one of the main products in Iran, the USA, and Turkey [1, 2]. According to the Food and Agriculture Organization (FAO), pistachio production in 2018 was 551307 million tons, with Iran accounting for about 38% of world pistachio production in that year. According to statistics released by FAO, Iran is the largest producer of pistachios in the world [3]. An important part of pistachios is consumed as a snack. However, pistachio nuts are used to make cakes, ice cream, spreads, pistachio butter, and parts of kielbasa formulations. This product is also used as the main component in the preparation of several traditional Iranian cookies and sweets such as Gaz[1], Baklava[2], and Qottab[3] [4, 5].

Characterizing the gravitational and geometric properties of nuts is important in the design of harvest and post-harvest, transportation, and processing equipment, and calculation of storage silo capacity and other requirements. For example, properties such as bulk density, repository angle, and grain dimensions are essential for the design of storage spaces and material handling equipment [6, 7]. Koushki and Mokhtarian (2021) investigated the effect of different moisture levels on changes in gravitational and geometric properties of two important Iranian pistachio cultivars (Ahmad Aghaei and Badami Sefid). The results showed that the gravitational and geometrical indices of Ahmad Aghaei pistachio cultivar were higher compared to those of the Badami Sefid cultivar. In addition, a comparison of various models (Mohsenin, McCabe, and Jain and Ball models) used to evaluate the sphericity coefficient indicated McCabe and Jain and Ball models have the highest efficiency in predicting the sphericity coefficients of Ahmad Aghaei (R^2=0.9838) and Badami Sefid cultivars (R^2=0.878), respectively [1]. Other researchers have also conducted comprehensive research into the gravitational and geometrical properties of pistachio nuts and kernels [8, 9, 10, 11].

Extensive studies have addressed the use of machine learning systems (such as artificial neural networks or ANN, genetic algorithms or GA, fuzzy logic or FL, ant colony algorithms or ACA, etc.) to predict and optimize the parameters of drying kinetics and physical and qualitative properties of food and energy consumption during processing operations. In fact, these systems create complex nonlinear relationships between input and output variables and simultaneously examine several project variables to provide a robust judgment system for deciding and monitoring instantaneous process parameters online. Currently, the use of smart control systems has a special place as non-destructive tools in various industries (especially in the design of intelligent machinery for food processing and agricultural products). Some of the studies addressing the smartization of prediction of the physical properties of food and agricultural products (especially nuts) are reviewed here. Hosseinpour-Zarnaq et al. (2022) examined an acoustic signal-based deep learning approach for the smart sorting of pistachio nuts. Moreover, the results obtained from this approach were compared with other models (multilayer perceptron artificial neural networks and random forest algorithms). Their findings showed that the one-dimensional convolutional neural network model is more efficient in predicting online pistachio nut sorting systems than other networks [12]. Mohammadi-

Moghaddam and Razavi (2019) studied the application of ANNs and image processing (IP) to predict the gravitational properties (unit mass and density) of roasted pistachio nuts and kernels. Their results showed that the ANN was able to predict the unit mass of roasted pistachio nuts and kernels with a coefficient of determination (mean square error or MSE) of 0.946 (MSE=0.000201) and 0.904 (MSE=.000198), respectively. The data also revealed that the ANN did not have a high efficiency to predict pistachio density and the coefficients of determination for this parameter were less than 0.71 [13]. Nouri-Ahmadabadi et al. (2017) studied the design, development, and evaluation of an online grading system for peeled pistachios equipped with machine vision technology (MVT) and support vector machine (SVM). The SVM model with cubic polynomial function and 38 support vectors had the best accuracy (R^2=0.9917) in pistachio grading [14]. Omid (2011) in another study designed an expert system to sort splitting from non-splitting pistachio nuts through a decision tree and fuzzy logic (FL) classifier, which led to successful results and new achievements in applying intelligent systems in the design of pistachio processing machines [15].

Since intelligent systems have a special place in the food industry, this study aims to explore the feasibility of predicting the lateral area of pistachio (Ahmad Aghaei cultivar) using an ANN system to select the best type and configuration of neural network based on input variables (major, intermediate, and minor diameter). It should be noted that modeling and measuring the lateral surface of agricultural products in various processes such as drying, aeration, heating, and cooling is important.

2. Materials and methods

Sample preparation

In this study, pistachios (Ahmad Aghaei cultivar) were prepared from the local market and transferred to the laboratory in closed containers. Before conducting the experiments, pistachios were cleaned of waste and external residues. All experiments in this study were performed on pistachio kernels. The initial moisture content of pistachio samples was determined to be 4.33% (w.b). The initial moisture content of the samples was measured by placing the samples in an atmospheric oven at a temperature of 105°C for 48 h until a constant weight was reached [16, 17].

Measuring the physical parameters of pistachio kernels

In this study, the physical characteristics of pistachio kernels were tested at different moisture levels. To achieve the desired moisture levels, the required number of pistachio samples were placed in polyethylene bags, and the required amount of water was added to the sample to complete the conditioning operation according to the principles of mass balance and through Eq. (1) [18, 19]:

$$Q = \frac{W_i\,(X_f - X_i)}{(100 - X_f)} \tag{1}$$

Where Q is the amount of water added (kg), W_i is the mass of the sample (kg), X_i is the initial moisture content of the sample (% w.b.), and X_f is the final moisture content of the sample (% w.b.).

To achieve the intended moisture content of the samples and its uniform distribution throughout the pistachios, the samples were kept at 5±1°C for 7 days [20]. At the end of the conditioning process, to ensure uniform distribution of moisture in the samples and achieve the desired moisture, the moisture content of the samples was measured again [19].

Moreover, to determine the axial dimensions of the samples (major diameter or L, intermediate diameter or W, and minor diameter or T in millimeters), 100 pistachio samples were randomly selected and their dimensions were measured using a caliper (Vertex, M502, with an accuracy of 0.01 mm) at different levels of tested moisture [19]. The geometric mean diameter (D_g) of the sample was measured via Eq. (2):

$$D_g = \sqrt[3]{LWT} \qquad (2)$$

Where D_g is the geometric mean diameter (mm) and L, W, and T are the major, intermediate, and minor diameters, respectively [19]. The surface area of the samples was calculated by McCabe equation [19]:

$$S = \pi D_g^2 \qquad (3)$$

Where S is the surface area (mm^2) and π is a geometric number.

Mathematical modeling using the artificial neural network (ANN)

The artificial neural network was modeled using SPSS software (version 19) (2011). Two types of neural networks, i.e. multilayer perceptron (MLP) neural network and radial base function (RBF) neural network, were used for modeling in which the input layer consisted of three neurons (major diameter, intermediate diameter, and minor diameter) and the output layer consisted of one neuron (surface area or S). Thus, the ANN model was developed with 3 inputs and 1 output. The structure of the mentioned ANNs was optimized by examining the different network configurations and also evaluating the correlation between the neural network output and the experimental data. To optimize the MLP-neural network, various network parameters such as the neurons number (NN) in each hidden layer, type of activation function in the hidden layer and output, learning rate, and momentum coefficient were evaluated. To find the best configuration, several hidden layers with 2-42 neurons in each hidden layer, the learning rate of 0.4, momentum coefficient of 0.7, and sigmoid logarithm activation functions (Eq. 4) in the hidden layer and output were used [21].

$$\log sig(z) = (1+\exp(-z))^{-1} \qquad (0,+1) \qquad (4)$$

To model the neural networks, the data were first divided into two parts, so that 70% of the data were considered for training and the remaining 30% for network evaluation. To compare the efficiency of neural networks, the coefficient of determination (R^2) and mean relative error (MRE) were used (Eq. 5 and 6):

$$R^2 = 1 - \left[\frac{\sum_{i=1}^{N} (U_{p,i} - U_{e,i})^2}{\sum_{i=1}^{N} (\bar{U}_{p,i} - U_{p,i})^2} \right] \qquad (5)$$

$$MRE = \left(\frac{1}{N} \sum_{i=1}^{N} \frac{\left| U_{p,i} - U_{e,i} \right|}{U_{e,i}} \right) \times 100 \qquad (6)$$

Where $U_{e,i}$ is the experimental data in measurement i, $U_{p,i}$ is the data predicted by the ANN in measurement i, $\bar{U}_{p,i}$ is the average of the data predicted by the ANN in measurement i, and N is the number of observations [22, 23].

3. Results

In this study, a combination of different layers and neurons with a sigmoid logarithm activation function (in the hidden and output layers) was used to optimize the MLP and the RBF. The neural networks with a hidden layer, 2 to 42 neurons were randomly selected and the power of the network in predicting the surface area of pistachio kernels was estimated. To train the MLP, the error propagation learning algorithm with momentum was used in which the momentum coefficient for all networks was 0.7 and the learning rate was 0.4. To obtain a suitable number of training epochs, a single-layer experimental network was trained with 2 to 46 neurons in the hidden layer and different numbers of training epochs. The results indicated that a total of 7000 training epochs had the best accuracy in predicting the surface area of pistachio kernels and they can also avoid over-training of the network. Table 1 shows the results of the optimization of the MLP-model with different configurations. As can be seen, the MLP-neural network with a hidden layer and the 1-18-3 configurations, i.e. a network with 3 inputs, 18 nodes in the hidden layer, and 1 output had the best performance in predicting the surface area of pistachio kernels (MRE=0.1149; R^2=0.895).

The RBF-network is trained faster than a MLP-model. The assessment of the RBF-network showed that the 1-42-3 configuration, i.e. a network with 3 inputs, 42 processing elements (neurons) in the hidden layer, and 1 output was most efficient in predicting the surface area of pistachio kernels (MRE=0.014; R^2=0.954). In general, a comparison of the performance of neural network models showed that the neural network with a RBF has higher accuracy in predicting the surface area of pistachio kernels and was selected as the best neural network model.

Table 1. Comparing the performance of multilayer perceptron (MLP) and radial base function (RBF) neural networks in predicting the surface area of pistachio kernels with three input variables (L, W, and T).

Number of hidden layers	Number of neurons	Type of networks			
		Multilayer perceptron (MLP)		Radial base function (RBF)	
		R^2	MRE	R^2	MRE
1	2	0.451	0.8568	0.606	0.3979
1	6	0.671	0.3424	0.713	0.2873
1	10	0.516	0.5103	0.796	0.2024
1	14	0.291	0.9761	0.826	0.1707
1	18	0.895	0.1149	0.829	0.1590
1	22	0.642	0.3708	0.917	0.0805
1	26	0.658	0.3792	0.917	0.0738
1	30	0.826	0.1791	0.931	0.0652
1	34	0.301	0.8986	0.939	0.0587
1	38	0.682	0.4614	0.944	0.0527
1	42	0.837	0.1951	0.954	0.0314

4. Discussion

An ANN consists of a set of internally connected neurons that can estimate output solutions based on input data. Neural networks are usually layered and have a regular form. The first layer into which input information and data enter is called the input layer. The middle layers are the hidden layers and the last layer that provides the output solutions of the model is called the output layer [24]. The simplest and most common type of neural network used in many engineering sciences, including the present study, is the MLP-neural network with an observer that uses the error post-propagation method for training. In this network, the number of input layer neurons is equal to the number of input vector elements and the number of output layer neurons is equal to the number of output vector elements. The accurate and realistic analysis to find the number of middle layer neurons is a complicated task. However, it can be suggested that the number of middle layer neurons is a function of the input vector elements and also the maximum number of areas of the input space that are linearly separated [25]. As mentioned earlier, the process of modeling the surface area of pistachio kernels was analyzed by two neural network models (MLP and RBF). Finally, the radial base function (RBF) model was selected as the superior network due to the lowest error rate and highest coefficient of determination. Mokhtarian et al. (2020) examined the performance of the machine learning system to predict some engineering properties (surface area, volume, mass, and particle density) of almond kernels. Their findings indicated that the ANN-model had a better performance in predicting the engineering properties of almond kernels. In addition, they reported that the MLP-neural network than RBF-model had the best performance in predicting the surface area, and the value of the coefficient of determination for this physical index (i.e., the surface area) was higher than 0.903 [26].

5. Conclusion

In this study, the surface area of pistachio kernels was predicted using two neural network models (multilayer perceptron and radial base function) and then different models were compared to choose the best neural network model. Following the data in this study, the RBF-model was selected as the best neural network model in predicting the surface area of pistachio nuts. Thus, the neural network was more efficient in estimating the surface area of pistachio kernels with a coefficient of determination of 0.954, which confirms the high accuracy of the neural network model. Accordingly, future studies can employ other optimization tools such as fuzzy-neural systems as well as other ANN models networks and different activation functions, compare them with experimental models, and evaluate the power of these tools in estimating engineering properties of pistachio kernels and nuts.

Acknowledgement

The current manuscript is extracted from a research study approved by the Islamic Azad University, Roudehen Branch, Roudehen, Iran. The authors would like to express their gratitude to the Vice Chancellor for Research and Technology at Islamic Azad University of Roudehen for the laboratory supports.

Conflict of interest: The authors declared no conflict of interest.

References

1- Koushki F, Mokhtarian A. The effects of moisture content and type of cultivars on physical properties of pistachio nuts. Pistachio Health J. **2021**; 4(4):*91-107*.

2- Mokhtarian M, Kalbasi-Ashtari A, Tavakolipour H. Performance of fuzzy logic model to foresee of ground pistachio nut isotherm. Pistachio Health J. **2020**; 3(1):*18-31*.

3- FAO. FaoStat Database. Available from http://faostat.fao.org; 2018.

4- Mokhtarian M, Tavakolipour H. The effects of solar drying systems and packaging types on the quality indicators of dried pistachio nuts. Pistachio Health J. **2019**; 2(1):*71-82*.

5- Mokhtarian M, Tavakolipour H, Kalbasi-Ashtari A. Energy and exergy analysis in solar drying of pistachio with an air recycling system. Drying Technol. **2016**; 34(12):*1484-1500*.

6- McLean K.A. Drying and Storing Combinable Crops. Farming Press; **1989**.

7- Navarro S, Noyes R.T. The Mechanics & Physics of Modern Grain Aeration Management. CRC Press; **2002**.

8- Razavi S.M.A., Emadzadeh B, Rafe A, Mohammad Amini A. The physical properties of pistachio nut and its kernel as a function of moisture content and variety: Part I. Geometrical properties. Food Eng. **2007a**; 81:*209-217*.

9- Razavi S.M.A., Rafe A, Mohammadi Moghaddam T, Mohammad Amini A. Physical properties of pistachio nut and its kernel as a function of moisture content and variety. Part II. Gravimetrical properties. Food Eng. **2007b**; 81:*218-225*.

10- Kashaninejad M, Mortazavi A, Safekordi A, Tabil L.G. Some physical properties of Pistachio (*Pistacia vera* L.) nut and its kernel. Food Eng. **2006**; 72:*30-38*.

11- Polat P, Aydin C, Erol A.B. Some Physical and Mechanical Properties of Pistachio Nut. J. Agricul. Sci. **2007**; 13:*237-246*.

12- Hosseinpour-Zarnaq M, Omid M, Taheri-Garavand A, Nasiri A, Mahmoudi A. Acoustic signal-based deep learning approach for smart sorting of pistachio nuts. Postharvest Biology Technol. **2022**; 185:*111778*.

13- Mohammadi-Moghaddam T, Razavi S.M.A. Application of artificial neural networks (ANN) and image processing for prediction of gravimetrical properties of roasted pistachio nuts and kernels. J. Nuts. **2019**; 10(2):*117-125*.

14- Nouri-Ahmadabadi H, Omid M, Mohtasebi S.S, Soltani Firouz M. Design, development, and evaluation of an online grading system for peeled pistachios equipped with machine vision technology and support vector machine. Info. Proc. Agricul. **2017**; 4:*333-341*.

15- Omid M. Design of an expert system for sorting pistachio nuts through a decision tree and fuzzy logic classifier. Expert Sys. Applications. **2011**; 38(4):*4339-4347*.

16- Mokhtarian M, Tavakolipour H, Kalbasi-Ashtari A. Effects of solar drying along with air recycling system on physicochemical and sensory properties of dehydrated pistachio nuts. LWT-Food Sci. Technol. **2017**; 75:*202-209*.

17- Tavakolipour H, Mokhtarian M, Kalbasi-Ashtari A. Lipid oxidation kinetics of pistachio powder during different storage conditions. J. Food Proc. Eng. **2017**; 40e12423.

18- Kingsly A.R.P, Singh D.B, Manikantan M.R, Jain R.K. Moisture dependent physical properties of dried pomegranate seeds (Anardana). J. Food Eng. **2006**; 75:*492-496*.

19- Mokhtarian M, Tavakolipour H, Hamedi H, DaraeiGarmakhany A. Performance of machine learning system to the prediction of almond physical properties. Food Health J. **2020**; 3(4):*7-13*.

20- Mokhtarian M, Tavakolipour H, Kalbasi-Ashtari A, Koushki F. The effects of solar drying on drying kinetics and effective moisture diffusivity of pistachio nut. Food Health. **2021**; 4(4):*7-12*.

21- Tavakolipour H, Mokhtarian M. Neural network approaches for prediction of pistachio drying kinetics. Int. J. Food Eng. **2012**; 8(3). DOI: 10.1515/1556-3758.2481.

22- Mokhtarian M, Tavakolipour H. The application of the neural network method for the prediction of the

osmotic factors of crookneck squash. J. Food Biosci. Technol. **2016**; 6(2):*23-30.*
23- Mokhtarian M, Garmakhany A. Production of ultrasound osmotic dehydration properties of courgette by ANN. J Quality Assurance Safety Crops Foods. **2017**; 9(2):*161-169.*
24- Tavakolipour H, Mokhtarian M, Kalbasi-Ashtari A. Intelligent monitoring of zucchini drying process based on fuzzy expert engine and ANN. J Food Proc. Eng. **2014**; 37:*474-481.*
25- Ghodsvali A.R, Farzaneh V, Bakhshabadi H, Zare Z, Karami Z, Mokhtarian M, Carvalho I.S. Screening of the aerodynamic and biophysical properties of barley malt. Int. Agrophys. **2016**; 30:*457-464.*
26- Mokhtarian M, Tavakolipour H, Hamedi H, Daraei Garmakhany A. Performance of machine learning system to the prediction of almond physical properties. Food Health J. **2020**; 3(4):*7-13.*