The Pharmaceutical and Chemical Journal, 2017, 4(4):33-38

Available online www.tpcj.org

Research Article

ISSN: 2349-7092
CODEN(USA): PCJHBA

Application of artificial neural networks to authenticate Virgin Groundnut Oil adulterants using FTIR spectral data

Pandurangan M. K.[1], Murugesan S.[2], Gajivaradhan P.[1], Shettu N.[3]

[1]PG and Research Department of Mathematics, Pachaiyappa's College, Chennai, India
[2]PG and Research Department of Botany, Pachaiyappa's College, Chennai, India
[3]PG and Research Department of Zoology, Pachaiyappa's College, Chennai, India

Abstract The aim of the present study is to detect the presence of palm oil and rice bran oil in virgin groundnut oil by Artificial Neural Networks (ANN) model. The oil samples were analyzed by FT-IR spectroscopy. They were classified as virgin groundnut oil, and two types of adulterated edible oil samples, such as palm oil and rice bran oil in different proportions (i.e.5, 10 and 15%). On the basis of cross validation procedure, the best models were selected and were further used for blend samples for prediction. The results obtained show that the models clearly separate different groups and classify correctly the virgin groundnut oil from the palm and rice bran oils. Moreover a reasonable discrimination between both mixtures and virgin oil was achieved.

Keywords Virgin groundnut oil, palm oil, rice bran oil, FT-IR Spectroscopy

Introduction

Artificial Neural Networks (ANNs) were originally developed to mimic the learning process of the human brain and the knowledge storage functions. The sensitivity and specificity of neural network models have a better predictive power compared to logistic regression. Even when applied to an external dataset, the neural network models performed better than the logistic regression. Neural networks are beginning to play an important role in the daily life of all human activity, and several research areas have started gradually incorporating this technology.

Artificial neural networks (ANNs) have been applied in almost every aspect of food science over the past two decades, although most applications are in the development stage. ANNs are useful tools for food safety and quality analysis, which include modeling of microbial growth and from this predicting food safety, interpreting spectroscopic data, and predicting physical, chemical, functional and sensory properties of various food products during processing and distribution.

Of the standard ANNs, the MLP (Multi Layer Perception) is perhaps the most popular network architecture currently in use [1]. The present study indicates the better predictive capabilities of MLP neural network. An MLP model consists of an input layer, a hidden layer and an output layer. In the present study the MLP architecture had six input variables, four hidden layers with two hidden nodes and one output node. All the artificial neurons are arranged in a layered feed–forward topology. In the ANN model, an independent evaluation is using external validation data and the neural network models are performed well. The predicting capabilities of an MLP neural network showed good results and more applications were developed using the SPSS neural network program and JMP software, which can both run the MLP model [2].

ANNs are non-linear statistical data modeling tools. They can take into account outliers and nonlinear interactions among variables and can reveal previously unrecognized and/or weak relationships between governing input

variables and an outcome [3]. This is a notable strong point in the application of ANNs,which is hence construed as a powerful tool for detecting the aspects of adulteration, especially in the case of edible oils. And no doubt, the use of ANNs can be used to authenticate the adulteration of several entities.

ANN may be particularly useful when the primary goal is classification and is important when interactions or complex nonlinearities exist in the dataset [4]. ANN models are being increasingly used in the detection of adulterants in the edible oils. Though, several limitations to studies like the present one were noted by Anderson *et al*., (2011) [5], such as the lack of an ensemble approach and a priori variable selection, by eliminating these limitations, Andersson *et al*., (2011) [5], successfully constructed an ANN using six variables to predict adulterants of the virgin oils.

In the present study, the artificial neural network (ANN) model was used to detect the percentage of adulteration in virgin oil and adulterated oil. The results showed that, based on ANN as a pattern recognition technique, the spectral data cannot predict the percentage of adulteration in virgin oil, but can be used in the qualitative determination of adulteration in the edible oil. This mathematical model was also validated using an external validation set of blend samples (5–15%) and virgin samples. However, artificial neural networks could be trained successfully to distinguish virgin oils from those which had been adulterated.

Materials and Methods

In order to accomplish pattern recognition using ANN, virgin groundnut oil sample as well as the samples of virgin groundnut oil adulterated with the palm and rice bran oils in the varying proportions of 5, 10 and 15% were prepared and the FTIR spectroscopy was performed. The spectral data obtained were used to carry out the ANN. It was found that there is some main peak position of the virgin and adulterated oils in the FTIR spectrum chart. It is accomplished that oil recognition can take advantage of the difference in the relative intensities.

Then multilayer programme neural network has been performed which was trained by the chosen data from the spectrum database. In this manner the real experiment data were identified by MLP neural network. It was further observed that the FTIR spectrum of virgin oil and the adulterant oils was very different from each other, which was used to accomplish the pattern recognition of edible oils, under consideration. FTIR spectroscopy combined with MLP neural network was fast, high sensitivity, non-contact and offered a high recognition rate. Thus, the ANN proved to be a new technique to accomplish the edible oil pattern recognition and quality detection.

For the ANN classification, the number of hidden units and the weight decay need to be optimized; this can be done through a cross validation. In order to investigate the influence of the size of neurons in the hidden layer on the prediction accuracy, ANNs with neuron numbers ranging from 3 to 30 were developed with the weight decay fixed at 0.1.

The input layer for each probe comprises the assessed values of the variables. The nodes in the hidden layers are constructed by the programme. The value in the output layer for each individual is the ANN-predicted value. The predicted value is evaluated as correct or not, compared to the observed value. This information is used by the programme to adjust the weights between the nodes: when the prediction is correct, the weights of the connections are increased; and when the prediction is wrong, the weights of the connections are decreased. This process continues until all cases are processed, and is called training the network.

The ANN result is the most probable outcome for each case, and a quantitative estimate of how reliable this estimate is can be obtained with an accurate value between 0 and 1. This means that there is no directly comparable parameter, either on the sample level or the individual level, so that a common comparable quantitative characteristic has to be constructed.

Results and Discussion

The number of neurons in the present study is 4, based on the sum of square error convergence. In the present study, we find that our network is activated by training time. The case processing summary shows that 30 cases were assigned to the training sample, 15 cases to the test sample and that no cases were holdout and excluded from the analysis.

The final three-layer 6-2-1-feed-forward, back propagation ANN model with variables consisting of palm oil 5%, palm oil 10%, palm oil 15%, rice bran oil 5% rice bran oil 10% and rice bran oil 15% was developed and trained in 45 spectral data (Fig.1).

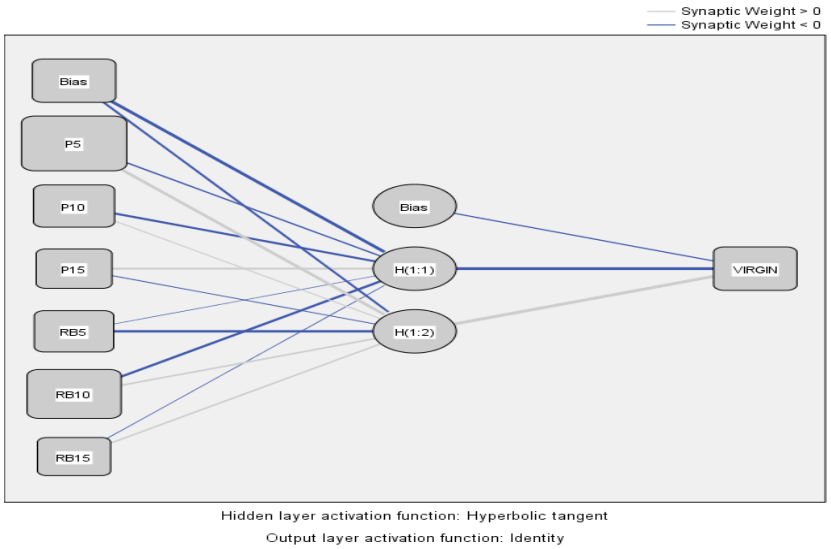

Figure 1: A neural network for the prediction of adulterants in virgin groundnut oil consisting of six inputs, variable (excluding –bias node), a hidden layer with three nodes (excluding –bias node) and one output variable.

The network information displays information about the neural network and it is useful for ensuring that the specifications are correct. The number of variables in the input layer is the number of covariates plus the total number of factor levels; a separate variable is created for each category of adulterants namely palm oil 5%, palm oil 10%, palm oil 15%, rice bran oil 5%, rice bran oil 10%, rice bran oil 15% and none of the categories are considered "redundant" units as is typical in modeling procedures. Likewise, a separate variable has been created for each category of virgin groundnut oil. Thus, the output layer will have only one variable, which is a pre-requisite for any modeling procedure. Covariates are rescaled using the adjusted normalized method. Automatic architecture selection has chosen 4 units in the hidden layer.

This structure is known as a feed forward architecture because the connections in the network flow forward from the input layer to the output layer without any feedback loops (Fig.1). The input layer contains the predictors with six variables (i.e.) covariates with one bias. The hidden layer contains unobservable nodes, or units. The value of each hidden unit is some function of the predictors; the exact form of the function depends, in part, upon the network type and, in part, upon user-controllable specifications. The improvement procedure has been discussed. The number of hidden layers is decided with the help of the least sum of square error . In the current model, the optimum level is reached with two hidden layers with a sum of square error values of 0.001. The output layer contains the responses. Since the history of default is a categorical variable with a single category, namely virgin groundnut oil, it is recoded as an indicator variable. Each output unit is some function of the hidden units. Again, the exact form of the function depends, in part, on the network type and, in part, on user-controllable specifications. The MLP (multiple layer perception) network allows a two hidden layer; in that case, each unit of the second hidden layer is a function of the units in the first hidden layer, and each response is a function of the units in the second hidden layer.

The error computations are based on the testing sample. The number of hidden units is determined by the testing data criterion. The best function of hidden units is the one that yields the smallest error in the data. The model summary (Table.1) displays information about the results of training, testing, and applying the final network to the holdout sample. Sum of squares error is an error function that the network tries to minimize during training and testing.

As shown in the normalized importance (Table.3; Fig.2) palm oil 5%, rice bran oil 10%, palm oil 10% and rice bran oil 5% were the most important predictors of persistent adulterants of sensitivity in order of priority. As previously mentioned, each neural network input element has an associated synaptic value, which is represented by a numerical

value that controls the input; the higher the synaptic value, the more relevant is the input for the result that is generated from the neural network. The figure.2 presents a relevance chart for each input field to obtain fitting results, information that becomes available after the creation of the node corresponding to the model generated. Thus, an adulteration of 5% palm oil with virgin groundnut oil does not reveal any significant form of adulteration in the groundnut oil, whereas; in the virgin oil an adulteration of rice bran oil at 10% reveal 74.5% of virgin groundnut oil. The other adulterants have shown significant changes in the virgin oil at more than 5% level and those adulterants are not explained further.

Table 1: Dependent variable

Model Summary

Training	Sum of Squares Error	.010
	Relative Error	.001
	Stopping Rule Used	1 consecutive step(s) with no decrease in error[a]
	Training Time	00:00:00.001
Testing	Sum of Squares Error	.004
	Relative Error	.001

Dependent Variable: Virgin Groundnut oil

Table 2: Hidden layer parameters

Parameter Estimates

Predictor		Predicted		
		Hidden Layer 1		Output Layer
		H(1:1)	H(1:2)	VIRGIN
Input Layer	(Bias)	-.771	-.362	
	P5	-.130	.642	
	P10	-.421	.088	
	P15	.371	-.071	
	RB5	-.001	-.460	
	RB10	-.427	.278	
	RB15	-.026	.173	
Hidden Layer 1	(Bias)			-.128
	H(1:1)			-.880
	H(1:2)			1.307

Table 3: Independent variable importance

	Importance	Normalized Importance
Palm oil 5%	0.312	100.0%
Palm oil 10%	0.141	45.1%
Palm oil 15%	0.103	32.9%
Rice bran oil 5%	0.129	41.3%
Rice bran oil 10%	0.232	74.5%
Rice bran oil 15%	0.084	26.8%

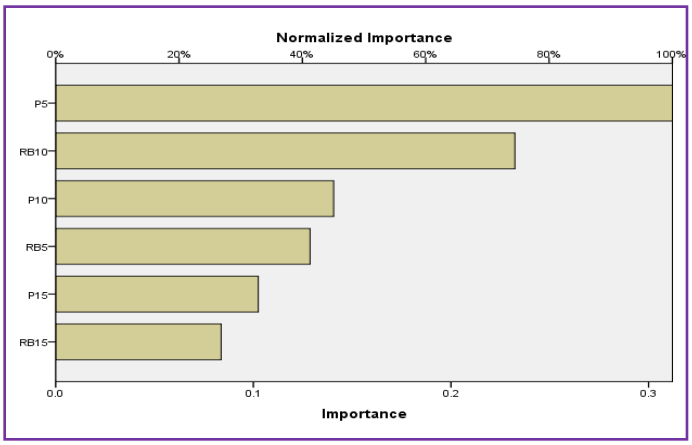

Figure 2: Sensitivity analysis of the input variables (The value shown for each input variable is a measure of its relative importance).

Table 4: Comparison of models

Size of Hidden Layer	Sum of Squares Error	Relative Error
2	0.017	0.003
3	0.006	0002
4	0.004	0.001

An ANN model with variables and bias was developed and trained in different spectral data and the results of this study demonstrate that the adulteration of 5% palm oil with virgin groundnut oil does not reveal any significant form of adulteration in the groundnut oil, whereas, in the virgin oil, an adulteration of rice bran oil at 10% reveal 74.5% of virgin groundnut oil.

Neural networks are beginning to play an important role in the daily life of all human activity, and thus industrial operations are gradually incorporating this technology into their work routines. Due to the robustness and efficacy of ANNs to solve complex problems, these methods have been widely employed in several research fields such as medicinal chemistry, pharmaceutical research, theoretical and computational chemistry, analytical chemistry, biochemistry, food research, etc. Therefore, ANN techniques can be considered as valuable tools to understand the main mechanisms involved in chemical problems.

Nowadays, the evolution of computer science (software and hardware) has allowed the development of many computational methods used to understand and simulate the behavior of complex systems. In this way, the integration of technological and scientific innovation has helped the treatment of large databases of chemical compounds in order to identify possible patterns. However, people who can use computational techniques must be prepared to understand the limits of applicability of any computational method and to distinguish between those opportunities which are appropriate to apply ANN methodologies to solve chemical problems.

The field of neural networks is very diverse and opportunities for future research exist in many aspects, including data pre-processing and presentation, architecture selection, and application [6-7]. The next logical step for the research is to improve further the performance of Neural Networks, for this application, perhaps through better training methods, better architecture selection, or better input.

Conclusion

In conclusion, an artificial neural network model with variables consisting of various proportions of adulterants may be useful for predicting the detection of virgin edible oil adulterants.By analyzing results that are obtained by the neural network with two, three and four hidden layer neurons, it is concluded that the best neural network configuration is probably four hidden neurons, which presents both a higher correlation and decreased error between actual data and estimated values. The goal is to be parsimonious, combining a good fit with as few neurons as possible.

Conflicts of interest

None declared.

References

1. Saftoiu, A., Vilmann, P., Gorunescu, F., Janssen, J., Hocke, M and Larsen, M. (2012). Efficacy of an artificial neural network-based approach to endoscopic ultrasound elastography in diagnosis of focal pancreatic masses. *Clin Gastroenterol epatol*. 10(1):84–90.
2. Hong, W.D., Ji, Y.F., Wang, D., Chen, T.Z and Zhu, Q.H. (2011). Use of artificial neural network to predict esophageal varices in patients with HBV related cirrhosis. *Hepat Mon1*. 1: 544-7.
3. Sall, J., Creighton, L., Lehman, A., Cary, N.C. SAS Pub. (2007). *JMP start statistics a guide to statistics and data analysis using JMP*. Safari Tech Books Online. SAS Press series. 4th ed.
4. Ji Zhao, Wei Chen Wenbo Xu. (2009). Non-linear system identification based on Radial Basis Function Neural Network Using Improved Particle Swarm Optimization". Fifth International conference on National Computation (409-413).
5. Anderson, B., Anderson, R., Ohlsson, M and Nilsson, J. (2011). Prediction of severe acute pancreatitis at admission to hospital using artificial neural networks. *Pancreatology*.11 (3):328–35.
6. Azoff, E.M. (1994). Neural network time series forecasting of financial market. John Wiley & Sons Ltd.
7. Swales, G.S and Yoon, Y. (1997). Applying artificial neural networks to investment analysis. *Financial Analysts Journal*. 48:78-80.