Artificial neural networks in foodstuffs: a critical review

S. Goyal*, G.K. Goyal
National Dairy Research Institute, Karnal -132001, India.

*Corresponding author: National Dairy Research Institute, Karnal -132001, India.

ARTICLE INFO

Article history:
Received 28 September 2012
Accepted 18 October 2012
Available online 30 October 2012

Keywords:
Artificial neural networks
Soft computing
Artificial intelligence
Foodstuffs
Review

ABSTRACT

This paper provides a critical review of literature concerning the artificial neural networks (ANN) in foodstuffs. The main aim is to provide background information, motivation for applications and an exposition to the methodologies employed in the development of ANN techniques in foodstuffs. This review includes that all the latest works on the application of ANN to foodstuffs which have been reported excellently with positive and encouraging results. This review paper highlights the methodologies and algorithms employed for ANN models suitable for various foodstuffs, viz., avocados, tomatoes, cherries, grape, mosambi juice, apple juice, chicken nuggets, pistachio nuts, potato chips, kalakand, cakes, processed cheese, butter, milk and other foodstuffs. This review paper would be very beneficial for those working in food industry, academicians, students, researchers, scientists, factories manufacturing the food products and regulatory authorities, as it provides comprehensive latest information.

© 2012 Sjournals. All rights reserved.

1. Introduction

Artificial Neural Networks (ANN) is authoritative modeling techniques that exhibit analogies to the way arrays of neurons function in biological learning and memory. ANNs offer numerous benefits over conventional modeling.
techniques because they can model based on no assumptions concerning the nature of the phenomenological mechanisms and understanding the mathematical background of problem underlying the process, and the ability to learn linear and nonlinear relationships between variables directly from a set of examples. The fundamental building blocks of ANN are units called nodes (neurons) comparable to biological neurons and weighted connections that can be likened to synapses in biological systems. Nodes are simple information processing elements. A well trained ANN can be used as a predictive model for a specific application. The prediction by a well trained ANN is normally faster than the mathematical models. Several authors have shown greater performances of ANN as compared to regression models. An ANN model can predict multiple dependent variables based on multiple independent variables, where a mathematical model is only able to predict one dependent variable at a time (Haykin, 1994; Amin et al., 2009; Vinod and Vikrant, 2002; Park et al., 2005; Zhang et al., 2002). ANN modelling in foodstuffs has been useful and successfully applied. ANNs are vibrant new tools to evaluate food quality, analyze shelf life and predict various properties of foodstuffs. The published literature is presented in the next section.

2. Avocados

ANN and hyperspectral techniques were used to model quality changes in avocados during storage at different temperatures. Avocados were coated using a pectin-based emulsion and stored at different temperatures (10, 15, 20°C), along with uncoated control samples. At different time intervals during storage period, respiration rate, total color difference, texture and weight loss of samples were measured as conventional quality parameters. Hyperspectral imaging was used to evaluate spectral properties of avocados. Multilayer ANNs were used in two ways to develop models for predicting quality parameters during storage. In the first set, ANN models were developed based on principal components of hyperspectral data as well as storage temperature and time. The optimal configuration of neural network model was obtained by varying the different model parameters. Results indicated ANN models to be accurate and versatile and they predicted the quality changes in avocado fruits better than the conventional regression models. Furthermore, the storage time–temperature-based ANN models were better than the hyperspectra-based ANN models (Maftoonazad et al., 2011).

3. Tomatoes

Movagharnejad and Nikzad (2007) studied the experimental works on drying of tomatoes in a tray dryer covering different variables like power of heater and air flow velocity. The data were modeled using ANN and empirical mathematical equations, and the results were compared with experimental data, which showed that the predictions of the ANN model fit the experimental data more accurately in comparison to the various mathematical equations.

4. Cherries

Guyer and Yang (2000) created a machine vision system to identify different types of tissue characteristics on cherries. It consists of an enhanced NIR range vidicon black and white camera (sensing range 400–2000 nm), a monochromatic controlled light source, and a computer. Multiple spectral images of cherry samples were collected over the 680–1280 nm range at increments of 40 nm. Using the spectral signatures of different tissues on cherry images, ANNs were applied to pixel-wise classification. An enhanced genetic algorithm was applied to design the topology and evolve the weights for multi-layer feedforward ANN. An average of 73% classification accuracy was achieved for correct identification as well as quantification of all types of cherry defects. No false positives or false negatives occurred; errors resulted only from misclassification of defect type or quantification of defect.

5. Grape

Multilayer perceptron ANN and radial basis function networks predicted ochratoxin A concentration over time in grape-based cultures of aspergillus carbonarius under different conditions of temperature, water activity
(a<sub>n</sub>) and sub-inhibitory doses of the fungicide carbendazim. The results obtained established ANNs as useful tools that should be fully explored in the field of food safety (Mateo et al., 2009).

6. Mosambi juice

ANNs described the permeate flux and permeate concentration (total soluble solid) profiles during the ultrafiltration of synthetic fruit juice and mosambi juice dynamically. It aimed to predict the permeate flux and total soluble solid of the permeate as a function of transmembrane pressure, sucrose, pectin concentration in the feed and the processing time. A multilayer feedforward ANN structure with input, output and hidden layers was used in this study. The backpropagation algorithm was used for training of ANN models. Two ANN models were constructed to predict the permeate flux and the total soluble solids in the permeate using the filtration data of the synthetic juice. The modeling results showed that there was an agreement between the experimental data and predicted values, with mean absolute errors (MAE) less than 1% of the experimental data. Also the trained networks were able to accurately capture the non-linear dynamics of synthetic fruit juice and the actual mosambi juice even for a new condition that had not been used in the training process (Rai et al., 2005).

7. Apple juice

Freeze-drying of foamed and non foamed apple juice was studied in order to assess if there was a reduction in process time due to foaming. Foams were prepared by whipping apple juice with methylcellulose or egg albumin at different concentrations. Foamed and non foamed juice samples having different thickness and different initial weight were frozen at – 40º C, and then freeze dried at 20º C during 48 h under vacuum. Sample weight loss and temperature were followed at different process times. A mathematical model based on ANN was developed to represent foam kinetics and temperature curves during freeze-drying. The observations revealed that freeze-drying of foamed materials is limited by heat transfer, while for non foamed ones, by mass transfer, and the insulation property characteristic of foams was more significant in slowing down the freeze-drying process than the increased surface area available for mass transfer due to foaming. From the study it was concluded that the ANN can be used to obtain excellent predictions of moisture content and temperature during the freeze-drying process (Raharitsifa and Ratti, 2010).

8. Chicken nuggets

A non-destructive, image-based method was evaluated for predicting mechanical properties of fried, breaded chicken nuggets. The textural parameters of interest, namely maximum load, energy to break point, and toughness of fried chicken nuggets were measured. Values of the parameters changed over frying time. Images of the chicken nuggets were collected at different frying stages and five image texture indices were extracted using co-occurrence matrix. A multiple-layer feed-forward neural network was established to predict the three mechanical parameters. The correlation coefficients of the predicted results with those from mechanical tests were above 0.84 (Qiao et al., 2007).

9. Beef fillet

A series of partial least squares (PLS) models were employed to correlate spectral data from FTIR analysis with beef fillet spoilage during aerobic storage at different temperatures (0, 5, 10, 15, and 20ºC). The performance of the PLS models was compared with a three-layer feedforward ANN developed using the same dataset. FTIR spectra were collected from the surface of meat samples in parallel with microbiological analyses to enumerate total viable counts. Sensory evaluation was based on a three-point Hedonic scale classifying meat samples as fresh, semi-fresh, and spoiled. The purpose of the modelling approach employed in this work was to classify beef samples in the respective quality class as well as to predict their total viable counts directly from TIR spectra. The results obtained demonstrated that both approaches showed good performance in discriminating meat samples in one of the three predefined sensory classes. The PLS classification models showed performances ranging from 72.0 to 98.2% using the training dataset, and from 63.1 to 94.7% using independent testing dataset. The ANN classification model performed equally well in discriminating meat samples, with correct classification rates from
98.2 to 100% and 63.1 to 73.7% in the train and test sessions, respectively. PLS and ANN approaches were also applied to create models for the prediction of microbial counts. The performance of these was based on graphical plots and statistical indices (bias factor, accuracy factor and RMSE) (Efthathios et al., 2011).

10. Meat

An ANN based predictive model for Leuconostoc mesenteroides growth in response to temperature, pH, sodium chloride and sodium nitrite developed by Garcia-Gimeno et al. (2005) was validated on vacuum packed, sliced, cooked meat products and applied to shelf life determination. Lag-time (Lag), growth rate (Gr), and maximum population density (yEnd) of L. mesenteroides, estimated by the ANN model were compared to those observed in vacuum-packed cooked ham, turkey breast meat, and chicken breast meat stored at 10.5°C, 13.5°C and 17.7°C, using bias and accuracy factors. The ANN model provided reliable estimates for the three kinetic parameters studied; with a bias factor of 1.09; 0.73 and 1.00 for Lag, Gr and yEnd, respectively, and an accuracy factor of 1.26; 1.58 and 1.13 for Lag, Gr and yEnd, respectively. From the three kinetic parameters obtained by the ANN model, commercial shelf life were estimated for each temperature and compared with the tasting panel evaluation. The commercial shelf life estimated microbiologically, i.e., times to reach 106.5 cfu/g was shorter than the period estimated using sensory methods (García-Gimeno et al., 2005; Zurera-Cosano et al., 2005).

11. Iranian Flat Bread

ANN based thermal conductivity (K) prediction model was developed for Iranian flat breads. The experimental data needed for developing the models were obtained from a pilot-scale set-up. Breads were made from three different cultivars of wheat and were baked in an eclectic oven at three different baking temperatures (232°C, 249°C and 260°C). A data set of 205 conditions was used for developing ANN and empirical models. To model K using ANN, 16 different MLP (multilayer perceptron) configurations ranging from one to two hidden layers of neurons were investigated and their prediction performances were evaluated. The (4-3-5-1)-MLP network, which is a network having two hidden layers, with three neurons in its first hidden layer and five neurons in its second hidden layer, had the best results in predicting the thermal conductivity of flat bread. For this network $R^2$, MRE, MAE and SE were 0.988, 0.6323, $1.66 \times 10^{-3}$, and $8.56 \times 10^{-5}$, respectively. Overall, ANN models (with $R^2 \geq 0.95$) performed superior than the empirical model (with $R^2 = 0.870$) (Omid et al., 2011).

12. Pistachio Nuts

Drying kinetics of pistachio nuts was simulated using a multilayer feed-forward neural network (MFNN). Experiments were performed at five drying air temperatures (ranging from 40 to 80°C) and four input air flow velocities (ranging from 0.5 to 2 m/s) with three replicates in a thin-layer dryer. Initial moisture content in all experiments was held at about 0.3 kg/kg d.b. To find the optimum model, various multilayer perceptron (MLP) topologies, having one and/or two hidden layers of neurons were investigated and their prediction performances were evaluated. The (3-8-5-1)-MLP, namely, a network having eight neurons in the first hidden layer and five neurons in the second hidden layer resulted in the best-suited model estimating the moisture content of the pistachio nuts at all drying runs. For this topology, $R^2$ and MSE values were 0.9989 and 4.20E-06, respectively. A comparative study among MFNN and empirical models was also carried out. Among the empirical models, the logarithmic model with MSE = 7.29E-6 and $R^2 = 0.9982$ gave better predictions than the others. However, the MFNN model performed better than the Lewis, Henderson and Pabis, two-term, and Page models and was marginally better than the logarithmic model (Omid et al., 2009).

13. Potato Chips

The ANN modeling approach was used to predict acrylamide formation and browning ratio (%) in potato chips as influenced by time*temperature covariants. A series of feedforward type network models with back-propagation training algorithm were developed. Among various network configurations, 4-5-3-2 configuration was found as the best performing network topology. Four neurons in the input layer reflected the asparagine concentration, glucose concentration, frying temperature, and frying time. The output layer had two neurons
representing acrylamide concentration and browning ratio of potato chips. The ANN modeling approach was shown to successfully predict acrylamide concentration ($R^2 = 0.992$) and browning ratio ($R^2 = 0.997$) of potato chips during frying at different temperatures in time-dependent manner for potatoes having different concentrations of asparagine and glucose, thus indicating that ANN modeling is a useful predictive tool which considers only the input and output variables rather than the complex chemistry (Serpen and Gökmen, 2007).

14. Corn

Campbell et al. (1999) conducted a study to investigate methods of improving a near-infrared transmittance spectroscopy (NITS) amylose calibration that could serve as a rapid, nondestructive alternative to traditional methods for determining amylose content in corn. Calibrations were developed using a set of genotypes possessing endosperm mutations in single- and double-mutant combinations ranging in starch-amylose content (SAC) from -8.5 to 76%, relative to a standard curve. The influence of three factors were examined including comparing calibrations made against SAC versus grain amylose content (GAC), developing calibrations using partial least squares (PLS) analysis versus artificial neural networking, and using all samples in the calibrations set versus using progressively narrower ranges of SAC or GAC in the calibration set. Grain samples were divided into calibration and validation sets for PLS analysis while samples used in ANN were assigned to a training set, test set, and validation set. Performance statistics of the validation sets that were considered were the $R^2$, the standard error of prediction (SEP), and the ratio of the standard deviation of amylose values to the SEP (RPD), which were used to compare all NITS models. The study revealed an NITS prediction model for SAC ($R^2 = 0.96$, SEP = 5.1%, RDP = 3.8) of similar precision to the best GAC model ($R^2 = 0.96$, SEP = 2.7%, RPD = 3.5). Narrowing the amylose range of the calibration set generally did not improve performance statistics except for PLS models for SAC in which a decrease in SEP values was observed. In one model, the SEP improved while $R^2$ and RPD remained constant ($R^2 = 0.94$, SEP = 4.2%, RPD = 2.8) when samples with SAC values <20% were removed from the calibration set. Although the NITS amylose calibrations in this study are of limited precision, they may be useful when a rough screening method is needed for SAC. For example, NITS may be useful to detect severe contamination during transport and storage of specialty grains or to aid breeders when selecting for amylose content from large numbers of grain samples [20].

15. Honey

Seventy samples of honey of different geographical and botanical origin were analysed with an electronic nose. The instrument equipped with 10 Metal Oxide Semiconductor Field Effect Transistors (MOSFET) and 12 Metal Oxide Semiconductor (MOS) sensors were used to generate a pattern of the volatile compounds present in the honey samples. The sensor responses were evaluated by Principal Component Analysis and ANN. Good results were obtained in the classification of honey samples by using a neural network model based on a multilayer perceptron that learned using a backpropagation algorithm. The methodology is simple, rapid and results suggest that the electronic nose could be a useful tool for characterization and control of honey (Benedetti et al., 2004).

16. Corn

Yang et al. (2000) developed a backpropagation ANN model that could distinguish young corn plants from weeds. Although only the colour indices associated with image pixels were used as inputs, it was assumed that the ANN model could develop the ability to use other information such as shapes, implicit in these data. The 756x504 pixel images were taken in the field and were then cropped to 100x100-pixel images depicting only one plant, either a corn plant or weeds. There were 40 images of corn and 40 of weeds. The ability of the ANNs to discriminate weeds from corn was then tested on 20 other images. A total of 80 images of corn plants and weeds were used for training purposes. For some ANNs, the success rate for classifying corn plants was as high as 100%, whereas the highest success rate for weed recognition was 80%. This is considered satisfactory, given the limited amount of training data and the computer hardware limitations, suggesting that an ANN-based weed recognition system can potentially be used in the precision spraying of herbicides in agricultural fields.

17. Kalakand
Cascade ANN models were developed for predicting the shelf life of Kalakand, which is desiccated milk based sweetmeat. The network was trained with 100 epochs and number of neurons in single and double hidden layers varied from 1 to 30. Cascade models with single hidden layer having four neurons gave the best outcome (MSE 0.000592818; RMSE: 0.024347850; $R^2$: 0.992884381). Cascade models with two hidden layers having twenty neurons in the first layer and twenty neurons in the second layer gave best fit (MSE 0.000988770; RMSE: 0.03144471; $R^2$: 0.988125331) for predicting the shelf life of the product (Goyal and Goyal, 2011a). Soft computing and Elman ANN techniques have also been applied for predictive modelling of kalakand (Goyal and Goyal, 2012a; 2012b).

18. Cakes

Cascade Neural Network (CNN) and Probabilistic Neural Network (PNN) models were developed for shelf life detection of cakes. Input variables were the quality data of the product relating to moisture, titratable acidity, free fatty acids, peroxide value, and tyrosine; and overall acceptability sensory score assigned by the expert panel was the output. Mean Square Error (MSE), RMSE, Coefficient of Determination ($R^2$) and Nash - Sutcliffo Coefficient ($E^2$) were used in order to compare the prediction performance of the developed models. The best results of all the models were compared with each other, indicating that CNN model with single hidden layer having twenty five neurons was most suitable for shelf life detection of cakes (Goyal and Goyal, 2011b). In order to predict the shelf life of soft cakes Elman and self organizing simulated neural network models were proposed. In this study, the experimental data of the product, viz., moisture, titratable acidity, free fatty acids, tyrosine, and peroxide value were taken as input variables, and the overall acceptability sensory score as the output. Neurons in each hidden layers varied from 1 to 30. The network was trained with single as well as double hidden layers with 1500 epochs, and transfer function for hidden layer was tangent sigmoid; while for the output layer, it was pure linear function. The experiments revealed that the developed ANN models predicted the shelf life of soft cakes exceedingly well (Goyal and Goyal, 2011c). Other ANN modelling algorithms have also been used for predictive modelling of cakes (Goyal and Goyal, 2012c; 2011d).

19. Butter

The seasonal variations of the fatty acids composition of butters were investigated over three seasons during a 12-month study in the protected designation of origin Parmigiano-Reggiano cheese area. Fatty acids were analyzed by GC-FID, and then computed by ANN. Compared with spring and winter, butter manufactured from summer milk creams showed an optimal saturated/un-saturated fatty acids ratio (~8.89 and ~5.79%), lower levels of saturated fatty acids (~2.63 and ~1.68%) and higher levels of mono-unsaturated (~5.50 and ~3.45%), polyunsaturated fatty acids (~0.65 and ~0.17%), and rumenic acid (~0.55 and ~3.41%), while vaccenic acid had lower levels in spring and higher in winter (~2.94 and ~2.91%). ANN models were able to predict the season of production of milk creams, and classify butters obtained from spring and summer milk creams on the basis of the type of feeding regimens (Gori et al., 2011).

20. Processed cheese

Linear Layer (Train) and Generalized Regression ANN models have been reported for predicting the shelf life of processed cheese stored at 7-8º C. The comparison of the two developed models showed that Generalized Regression model with spread constant as 10 got best simulated with less than 1% RMSE. The study showed that computational intelligence models are quite effective in predicting the shelf life of processed cheese (Goyal and Goyal, 2012d). Several other ANN techniques have also been reported for modelling of processed cheese (Goyal and Goyal, 2012e; 2011e; 2012f; 2012g; 2012h; 2012i; 2012j).

21. Milk

The accuracy of milk production forecasts on dairy farms using a ffann (Feedforward ANN) with polynomial post-processing has been implemented. Historical milk production data was used to derive models that are able to predict milk production from farm inputs, using a standard ffann, a ffann with polynomial post-processing and
multiple linear regressions. Forecasts obtained from the models were then compared with each other. Within the scope of the available data, it was found that the standard ffann did not improve on the multiple regression technique, but the ffann with polynomial post processing did (Sanzogni and Kerr, 2001).

22. Burfi

Radial basis (exact fit) model was proposed for estimating the shelf life of an extremely popular milk based sweetmeat namely burfi. The input variables were the data of the product relating to moisture, titratable acidity, free fatty acids, tyrosine, and peroxide value; and the overall acceptability score was output. MSE, RMSE, $R^2$ and $E^2$ were applied for comparing the prediction ability of the developed models. The observations indicated exceedingly well correlation between the actual data and predicted values, with a high $R^2$ and $E^2$ establishing that the models were able to analyze non-linear multivariate data with very good performance and shorter calculation time. From the study it was concluded that the developed model, which is very convenient, less expensive and fast, can be a good alternative to expensive, time consuming and cumbersome laboratory testing method for estimating the shelf life of the product (Goyal and Goyal, 2012k). Some other neural network modelling techniques have also been reported for burfi (Goyal and Goyal, 2012l; 2012m).

23. Instant coffee drink

For predicting the shelf life of instant coffee drink, radial basis artificial neural engineering and multiple linear regression (MLR) models were developed. Colour and appearance, flavour, viscosity and sediment were taken as input variables; while overall acceptability sensory score was taken as output variable. The investigation revealed that MLR model was superior to radial basis model for predicting the shelf life of instant coffee drink (Goyal and Goyal, 2011f). A number of other ANN models have also been reported for coffee drinks (Goyal and Goyal, 2011g; 2011h).

24. Conclusion

In this paper, we have presented the latest review of ANN techniques focusing on foodstuffs. This paper would be beneficial to those who are working in food industry, academicians, students, researchers, scientists, factories manufacturing the food products and regulatory authorities. The interface between foodstuffs and ANN techniques is still blur. There is, therefore, a need to apply ANN techniques more vigorously to other foodstuffs for predictive modelling, as it is simple and cost effective. Since the computer has become a vital tool in modern day society, it becomes more important to have a computational approach within which various foodstuffs could be explored for predictive analysis. ANN modelling approach to foodstuffs has the potential of predicting food products shelf life, processing and physico-chemical properties.

References


Goyal, Sumit, Goyal, G.K., 2011e. Performance of generalized regression, radial basis (fewer neurons), and linear layer (design) computational ANN techniques for shelf life prediction of processed cheese. Int. J. A.I. Knowledge Discovery. 1(4), 12-15.


