ALGORITHM COMPUTING APPROACH OF GENERALIZED REGRESSION NEURAL NETWORK FOR PREDICTING SHELF LIFE OF PROCESSED CHEESE

SUMIT GOYAL1*, GYANENDRA KUMAR GOYAL1

1National Dairy Research Institute, Karnal, India

Abstract: The aim of this research is to develop Generalized Regression Artificial Neural Network (ANN) models for predicting shelf life of processed cheese. Processed cheese is protein rich food, and is a comparable supplement to meat protein. Mean Square Error, Root Mean Square Error, Coefficient of Determination and Nash - Sutcliffe Coefficient were used in order to compare the prediction ability of the developed models. The modeling results showed that there was exceptional agreement between the experimental data and the predicted values. The model might be an alternative method to control the expiration date of processed cheese.

Keywords: artificial intelligence, generalized regression, prediction, shelf life

1. INTRODUCTION

Since past many decades the technology for the preparation of cheese has improved significantly, and several value added products have appeared, out of which processed cheese which is generally prepared from cow’s milk medium ripened Cheddar cheese, is the one that is extremely popular among the consumers due to its unique body and texture, aroma and flavour, and sensory attributes [1]. Processed cheese is protein rich food, and is a comparable supplement to meat protein. The purpose of this research is to develop ANN Generalized Regression models for predicting the shelf life of processed cheese stored at ambient temperature.

ANNs are inspired by the early models of sensory processing by the brain. An ANN can be created by simulating a network of model neurons in a computer. By applying algorithms that mimic the processes of real neurons, one can make the network ‘learn’ to solve many types of problems. A model neuron is referred to as a threshold unit. It receives input from a number of other units or external sources, weighs each input and adds them up. If the total input is above a threshold, the output of the unit is one; otherwise it is zero. Therefore, the output changes from 0 to 1 when the total weighted sum of inputs is equal to the threshold. The points in input space satisfying this condition define a so called hyperplane. In two dimensions, a hyperplane is a line, whereas in three dimensions, it is a normal plane. Points on one side of the hyperplane are classified as 0 and those on the other side as 1. Thus, a classification problem can be solved by a threshold unit if the two classes can be separated by a hyperplane [2].

The shelf life is the length of time till a product retains its natural taste and quality. In today’s highly competitive market consumers look for healthy and delightful food products. To attain good quality of food products, prediction of the shelf life is important [1, 3-5]. Predicting shelf life of processed cheese in laboratory is very time consuming and expensive. Therefore, it is relevant to develop such a model that predicts shelf life of processed cheese at low cost and in less time.

* Corresponding author, email: thesumitgoyal@gmail.com and gkg5878@yahoo.com
© 2013 Alma Mater Publishing House
A rapid and non-destructive method has been devised for the characterization of chocolate samples based on diffuse reflectance near-infrared Fourier transform spectroscopy (DRIFTS) and ANNs [6]. 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 [7]. Ni and Gunasekaran observed that a three-layer ANN model is able to predict more accurately than regression equations for the rheological properties of Swiss type cheeses on the basis of their composition [8]. Jimenez-Marquez et al. were of the opinion that prediction of moisture in cheese of commercial production using neural networks models can be used both for research to develop the base of knowledge on production variables and their complex interactions, as well as for the prediction of cheese moisture [9]. A methodology was developed for time series sales forecasting for short shelf-life food products based on ANN models and evolutionary computing. The methodology was claimed to be particularly useful for manufacturers of fresh milk, since successful sales forecasting reduces considerably the lost sales and products returns [10]. Zhang et al. developed ANN models for rough rice drying to predict six performance indices [11]. ANNs have been effectively implemented for predicting shelf-life of soya milk [12]; processed cheese [4, 5, 13]; kalakand [14, 15]; burfi [16-19]; coffee drink [20]; food processing [21] and noodles [22].

2. MATERIAL AND METHODS

Generalized regression neural network models are a kind of radial basis network that is used for function approximation. Syntax: net = newgrnn (P,T,spread)

net = newgrnn (P,T,spread) takes three inputs,
P: R-by-Q matrix of Q input vectors
T: S-by-Q matrix of Q target class vectors
Spread: Spread of radial basis functions (default = 1.0) and returns is a new generalized regression neural network. The larger the spread, the smoother is the function approximation. To fit data very closely, a smaller spread is used than the typical distance between input vectors. To fit the data more smoothly, a larger spread is used. Newgrnn creates a two-layer network. The first layer has radbas neurons, and calculates weighted inputs with dist and net input with netprod. The second layer has purelin neurons, calculates weighted input with normprod, and net inputs with netsum. Only the first layer has biases. newgrnn sets, the first layer weights to P’, and the first layer biases are all set to 0.8326/spread, resulting in radial basis functions that cross 0.5 at weighted inputs of +/- spread. The second layer weights W2 are set to T [23, 24].

For development of the models data consisted of 36 observations of which 30 were used for training the neural network, and 6 for validation. Soluble nitrogen, pH, standard plate count, yeast & mould count, and spore count were taken as input parameters, and sensory score as output parameter (Figure 1).

![Fig. 1. Inputs and output parameters for ANN models.](image_url)

Many different combinations were tried and tested, as there is no defined rule of getting good results rather than hit and trial method. As the number of neurons increased, the training time also increased. The Neural Network Toolbox under MATLAB software was used for developing the models. Training pattern of models is presented...
in Figure 2. Mean Square Error: MSE (1), Root Mean Square Error: RMSE (2), Coefficient of Determination: R² (3) and Nash - Sutcliffe Coefficient: E² (4) were used in order to compare the prediction ability of the developed models, where \( Q_{\text{exp}} \) = Observed value; \( Q_{\text{cal}} \) = Predicted value; \( \bar{Q}_{\text{exp}} \) = Mean predicted value; \( n \) = Number of observations in dataset.

\[
MSE = \left[ \frac{1}{n} \sum_{i=1}^{N} \left( \frac{Q_{\text{exp}} - Q_{\text{cal}}}{Q_{\text{exp}}} \right)^2 \right]
\]  
(1)

Fig. 2. Training pattern of ANN models.
3. RESULTS AND DISCUSSION

Generalized Regression ANN model, as depicted in Table 1 was developed for predicting the shelf life of processed cheese.

Table 1. Results of Generalized Regression ANN model.

<table>
<thead>
<tr>
<th>Spread Constant</th>
<th>Neurons</th>
<th>MSE</th>
<th>RMSE</th>
<th>$R^2$</th>
<th>$E^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>30</td>
<td>0.000410575</td>
<td>0.020262643</td>
<td>0.979737357</td>
<td>0.999589425</td>
</tr>
<tr>
<td>20</td>
<td>30</td>
<td>0.00028398</td>
<td>0.016851716</td>
<td>0.983148284</td>
<td>0.99971602</td>
</tr>
<tr>
<td>30</td>
<td>30</td>
<td>0.000251274</td>
<td>0.015851629</td>
<td>0.984148371</td>
<td>0.999748726</td>
</tr>
<tr>
<td>40</td>
<td>30</td>
<td>0.000239544</td>
<td>0.0154772</td>
<td>0.9845228</td>
<td>0.999760456</td>
</tr>
<tr>
<td>50</td>
<td>30</td>
<td>0.000234042</td>
<td>0.01529844</td>
<td>0.98470156</td>
<td>0.999765958</td>
</tr>
<tr>
<td>60</td>
<td>30</td>
<td>0.000231095</td>
<td>0.015201814</td>
<td>0.984798186</td>
<td>0.999768905</td>
</tr>
<tr>
<td>70</td>
<td>30</td>
<td>0.000229263</td>
<td>0.015141422</td>
<td>0.984858578</td>
<td>0.999770737</td>
</tr>
<tr>
<td>80</td>
<td>30</td>
<td>0.000228094</td>
<td>0.015102771</td>
<td>0.984897229</td>
<td>0.999771906</td>
</tr>
<tr>
<td>90</td>
<td>30</td>
<td>0.000227292</td>
<td>0.015076199</td>
<td>0.984923801</td>
<td>0.999772708</td>
</tr>
<tr>
<td>100</td>
<td>30</td>
<td>0.000226782</td>
<td>0.015059289</td>
<td>0.984940711</td>
<td>0.999773218</td>
</tr>
</tbody>
</table>

The comparison of Actual Sensory Score (ASS) and Predicted Sensory Score (PSS) for ANN model is illustrated in Figure 3.

Fig. 3. Comparison of ASS and PSS.
It is observed from Table 1 that ANN model having 100 as spread constant and 30 neurons gave the best results, MSE = 0.00226782, RMSE = 0.015059289, $R^2 = 0.984940711$, $E^2 = 0.99773218$ suggesting that the model is quite efficient in predicting the shelf life of processed cheese stored at ambient temperature. Goyal and Goyal [24] developed Linear Layer (Train) and Generalized Regression models for predicting the shelf life of processed cheese stored at 7-8°C. Soluble nitrogen, pH, Standard plate count, Yeast & mould count, Spore count were taken as input variables, and the experimental sensory score was taken as output variable for developing models. The data consisted of 36 experimental observations, which were sub-divided into two sets, i.e., 30 (80% of observations) for training, and 6 (20% of observations) for validation. On comparing the two developed models, viz., Linear Layer (Train) ($MSE = 0.000115628, RMSE = 0.010753058, R^2 = 0.989246942, E^2 = 0.999884372$), and Generalized Regression ($MSE = 1.22701E-05, RMSE = 0.003502871, R^2 = 0.996497129, E^2 = 0.99998773$), it was observed that Generalized Model performed better than Linear Layer (Train) model. This study also confirms that Generalized Regression models are capable for predicting the shelf life of processed cheese.

4. CONCLUSION

Generalized regression ANN model was developed for predicting the shelf life of processed cheese. Soluble nitrogen, pH, standard plate count, yeast & mould count, and spore count were input parameters, and sensory score was the output parameter. Mean Square Error, Root Mean Square Error, Coefficient of Determination and Nash - Sutcliffe Coefficient were used in order to compare the prediction ability of the developed model. It was observed that ANN model with spread constant as 100 having 30 neurons gave the best result. The modeling results showed that there was excellent agreement between the experimental data and predicted values, with a high determination coefficient ($R^2 = 0.984940711$) indicating that the Generalized regression model was able to analyze nonlinear multi-variant data with very good performance, fewer parameters, and shorter calculation time. Considering the complexity of the sample, Generalized regression ANN model was found to be reliable for predicting the shelf life of processed cheese stored at ambient temperature. These results are in agreement with the earlier findings of Goyal and Goyal [25, 26] like other reported studies, they suggested that ANN models could provide an effective means of recognizing the patterns in data and accurately analyze data for prediction. Future work could include using other neural network algorithms and parameters for developing ANN models for predicting shelf life of processed cheese.

REFERENCES