# Cascade and Feedforward Backpropagation Artificial Neural Network Models For Prediction of Sensory Quality of Instant Coffee Flavoured Sterilized Drink

Sumit Goyal and Gyandera Kumar Goyal

**Abstract —** *Today flavoured milks have become very popular and they contain nutrients as compared with soft drinks. Flavoured milks are milks to which flavours have been added. Sterilized milk is the product made by heating to high temperature. The term "coffee" comes from Arabic "Qahwah" through the Turkish 'Kahvah' and was originally one of the names employed for wines in Arabic. Coffee as beverage is prepared from the roasted seeds (beans) of the coffee plant. Coffee is the second most important product in the international market in terms of volume trade and the most important in terms of value. Cascade forward and feedforward backpropagation artificial intelligence models for prediction of sensory quality of instant coffee flavoured sterilized drink were developed. Colour and appearance, flavour, viscosity and sediment were taken as input parameters. The Overall acceptability was used as output parameter for developing the artificial intelligence models. The dataset consisted of experimentally obtained 50 observations. The dataset was randomly divided into two disjoint subsets, namely, training set containing 40 observations (80% of total observations) and test set comprising of 10 observations (20% of total observations).Different combinations of several internal parameters, i.e. data pre-processing, data partitioning, number of hidden layers, number of neurons in each hidden layer, transfer function, error goal, etc., along with backpropagation algorithm based on Levenberg–Marquardt mechanism as training function.. The network was trained with 100 epochs. The number of neurons in each hidden layer varied from 1 to 20. The results of cascade forward and feed forward artificial intelligent models were evaluated with three types of prediction performance measures viz., Root Mean Square Error, coefficient of determination $R^2$ and Mean Square Error. Feedforward backpropagation artificial intelligence model exhibited best results (3.70% RMSE; 0.998 $R^2$ ; 0.0013 MSE) , followed by cascade forward artificial intelligence model (5.36% RMSE; 0.996 $R^2$ ; 0.0028 MSE). From the study, it is concluded that the developed feedforward backpropagation artificial intelligence model is better than cascade forward artificial intelligence model.*

Sumit Goyal is Sr. Research Fellow at National Dairy Research Institute, Karnal 132001, India (e-mail: thesumitgoyal@gmail.com).

Gyandera Kumar Goyal is Emeritus Scientist at National Dairy Research Institute, Karnal 132001, India (e-mail: gkg5878@yahoo.com).

**KeyWords — Artificial Intelligence, Artificial Neural Networks, Neurocomputing, Soft Computing, Shelf Life Prediction.**

## I. INTRODUCTION

Today flavoured milks have become very popular and they contain nutrients as compared with soft drinks. They are manufactured by using natural and synthetic flavours, viz., chocolate, vanilla, strawberry, banana, raspberry and orange. Flavoured milks are milks to which flavours have been added. Sterilized milk is the product made by heating milk to high temperature (121 $^{o}$C) with 15 m holding time so that it remains fit for human consumption for at least 6 months at room temperature. The term "coffee" comes from Arabic "Qahwah" through the Turkish 'Kahvah' and was originally one of the names employed for wines in Arabic. It was first food, then a wine, a medicine and finally a beverage. Its use as a beverage dates back many centuries and was first mentioned in the literature of Rhazes, an Arabian physician in about 900 A.D. Coffee as beverage, is prepared from the roasted seeds (beans) of the coffee plant. Coffee plant is a tropical evergreen shrub of the genus "coffee" of the Madder family (Rubiaceae). Coffee is the second most important product in the international market in terms of volume trade and the most important in terms of value. Artificial Neural Network (ANN) is a neurobiologically inspired paradigm that emulates the functioning of the brain based on the way that neurons work, because they are recognized as the cellular elements responsible for the brain information processing [1]. ANN models can detect patterns that relate input variables to their corresponding outputs in complex biological systems for prediction [2].Methods for improving network performance include finding an optimum network architecture and appropriate number of training cycles, using different input combinations [3].Goñi [4] observed that ANN provides a simple and accurate prediction method. Goyal and Goyal [5] developed radial basis and regressions models for forecasting shelf life of instant coffee drink. Linear layer (design) and time - delay methods of intelligent computing expert system predicted shelf life of soft mouth melting milk cakes [6]. For achieving good quality and safety of food products detection of shelf life is important. Goyal and Goyal developed central nervous system based intelligent computer engineering system to detect shelf life of soft mouth melting milk cakes [7]. Shelf life is the recommendation of time that products can be stored, during which the defined quality of a specified proportion of the goods remains acceptable under expected (or specified) conditions of distribution, storage and display[8].Prediction of shelf life in laboratory is quiet time consuming and an expensive activity. Therefore, it would be relevant to develop a system that could predict shelf life of instant coffee flavoured sterilized drink at low cost and in less time. They would be useful to coffee shops owners and food researchers.

## II. METHOD MATERIAL

### A. Data Sampling

The dataset consisted of experimentally obtained 50 observations. The dataset was randomly divided into two disjoint subsets, namely, training set containing 40 observations (80% of total observations) and test set comprising of 10 observations (20% of total observations).

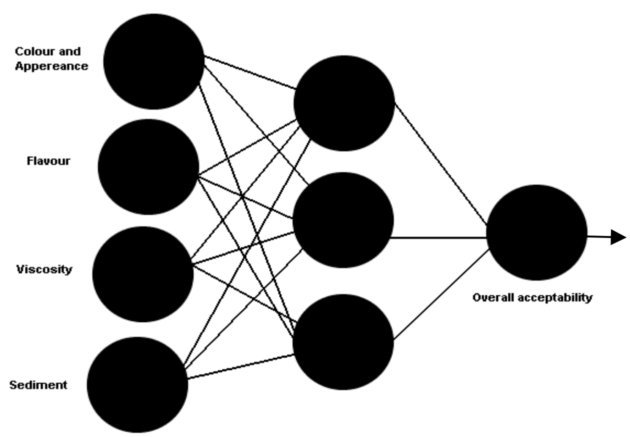



**Fig.1. Input and output parameters of artificial intelligence models**

Colour and appearance, flavour, viscosity and sediment were used as input parameters. The Overall acceptability was used as output parameter for developing the artificial intelligence models (Fig.1).

### B. Artificial Neural Networks Models

Cascade forward backpropagation and feedforward backpropagation artificial intelligence models were trained with the dataset of instant coffee flavoured sterilized drink. Different combinations of several internal parameters, *i.e.,* data preprocessing, data partitioning approach, number of hidden layers, number of neurons in each hidden layer, transfer function, error goal, *etc* were tried. Different variants of the backpropagation algorithm were explored like gradient descent algorithm with adaptive learning rate, Fletcher–Reeves update conjugate gradient algorithm, Polak–Ribiére Update conjugate gradient algorithm, Powell–Beale restarts conjugate gradient algorithm, BFG quasi-Newton algorithm, Levenberg–Marquardt algorithm and Bayesian regularization. Levenberg–Marquardt algorithm gave better results, hence it was used as training function; sum square error was performance function used during training of feedforward backpropagation neural network. There is no generalized method to determine the optimum values for number of hidden layers, neurons in each hidden layer, *etc.*, as they are function of expected intelligence. Neural network toolbox under Matlab 7.0 software was used for developing the artificial intelligence models.

### C. Feedforward backpropagation (FB) Model

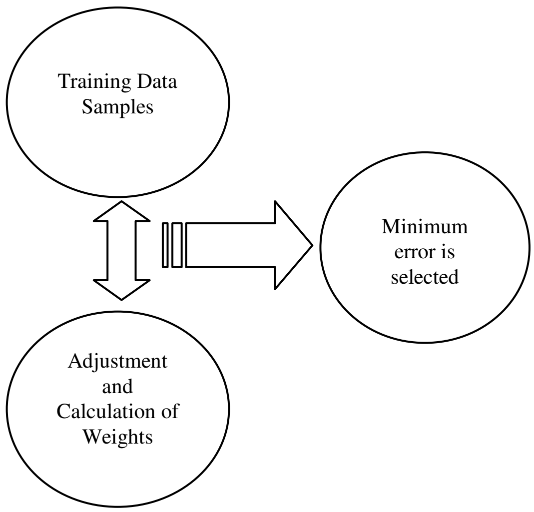
FB artificial intelligence model consists of input, hidden and output layers. Backpropagation learning algorithm was used for learning these networks. During training this network, calculations were carried out from input layer of network toward output layer, and error values were then propagated to prior layers. Feedforward networks often have one or more hidden layers of sigmoid neurons followed by an output layer of linear neurons. Multiple layers of neurons with nonlinear transfer functions allow the network to learn nonlinear and linear relationships between input and output vectors. The linear output layer lets the network produce values outside the range –1 to +1. On the other hand, outputs of a network such as between 0 and 1 are produced, then the output layer should use a sigmoid transfer function (logsig) [9].

### D. Cascade forward (CF) Model

CF models are similar to feed-forward networks, but include a weight connection from the input to each layer and from each layer to the successive layers. While two-layer feed-forward networks can potentially learn virtually any input-output relationship, feed-forward networks with more layers might learn complex relationships more quickly. The function newcf creates cascade-forward networks. For example, a three-layer network has connections from layer 1 to layer 2, layer 2 to layer 3, and layer 1 to layer 3. The three-layer network also has connections from the input to all three layers. The additional connections might improve the speed at which the network learns the desired relationship [9]. CF artificial intelligence model is similar to feedforward backpropagation neural network in using the backpropagation algorithm for weights updating, but the main symptom of this network is that each layer of neurons related to all previous layer of neurons [10].Tan-sigmoid transfer function, log - sigmoid transfer function and pure linear threshold functions were used to reach the optimized status[9].The performance of cascade forward backpropagation and feedforward backpropagation were evaluated using Root Mean Square Error (RMSE) Eq. (1), Mean Square Error (MSE) Eq. (2) and $R^2$ Eq. (3) technique.

The following criterion of root mean square error has defined to minimize the training error.

**Training Pattern of ANN's**




**Prediction performance measures**

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

$$RMSE = \sqrt{\frac{1}{n} \left[ \sum_{1}^{N} \left( \frac{Q_{\exp} - Q_{cal}}{Q_{\exp}} \right)^2 \right]} \quad (2)$$

$$R^2 = 1 - \left[ \sum_{1}^{N} \left( \frac{Q_{\exp} - Q_{cal}}{Q_{\exp}^{2}} \right)^2 \right] \quad (3)$$

Where, $Q_{\exp}$ = Observed value; (4) $Q_{cal}$ = Predicted value; $\overline{Q}_{\exp}$ =Mean predicted value; $n$ = Number of observations in dataset. MSE (1), RMSE (2) and $R^2$ (3) were used in order to compare the prediction performance of the developed models. The best score for $R^2$ measure is 1 and for other measures is zero.

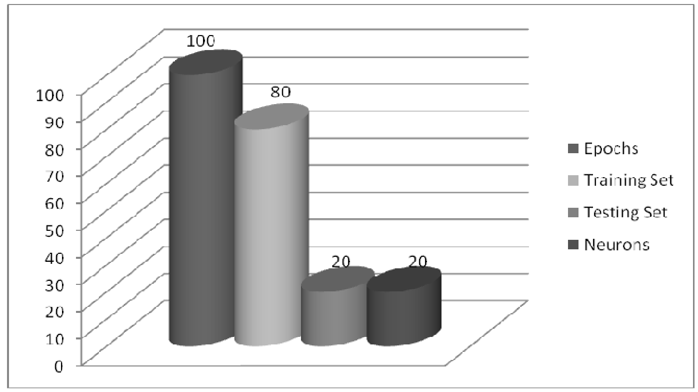



**Fig.2 Parameters used for artificial intelligence model**

The dataset were divided 80% for training, and 20% for testing. The number of neurons in each hidden layer varied from 1 to 20 as shown in Fig. 2. Weights and biases were randomly initialized. The network was trained with up to 100 epochs. An artificial intelligence model consists of a large number of processing elements called neurons. Weights and biases were randomly initialized. Weight is information used by neural network to solve a problem. In this research, two artificial intelligence models have been developed by employing feedforward backpropagation and cascade forward techniques.

## III. RESULTS AND DISCUSSION

FB artificial intelligence model performance matrices for predicting sensory quality are presented in Table 1.

**Table 1: Performance of FB artificial intelligence model for predicting sensory score**

| Neurons | MSE | RMSE | $R^2$ |
|---|---|---|---|
| 6 | 16.67 | 0.959 | 0.0277 |
| 8 | 8.85 | 0.991 | 0.0073 |
| 19 | 16.52 | 0.960 | 0.0272 |
| 5 | 15.96 | 0.963 | 0.0254 |
| 10 | 6.44 | 0.995 | 0.0041 |
| 12 | 5.38 | 0.996 | 0.0028 |
| 13 | 13.77 | 0.974 | 0.0189 |
| 15 | 5.92 | 0.945 | 0.0035 |
| **17** | **3.70** | **0.998** | **0.0013** |
| 19 | 5.88 | 0.996 | 0.0034 |

CF artificial intelligence model performance matrices for predicting sensory quality are given in Table 2.

**Table 2: Performance of CF artificial intelligence model for predicting sensory score**

| Neurons | MSE | RMSE | $R^2$ |
|---|---|---|---|
| 4 | 19.80 | 0.937 | 0.0392 |
| 5 | 20.16 | 0.934 | 0.0406 |
| **7** | **5.36** | **0.996** | **0.0028** |

| | | | |
|---|---|---|---|
| 9 | 14.77 | 0.969 | 0.0218 |
| 10 | 6.97 | 0.995 | 0.0048 |
| 12 | 6.22 | 0.995 | 0.0038 |
| 15 | 20.25 | 0.924 | 0.0410 |
| 17 | 13.95 | 0.973 | 0.0194 |
| 18 | 16.09 | 0.962 | 0.0259 |

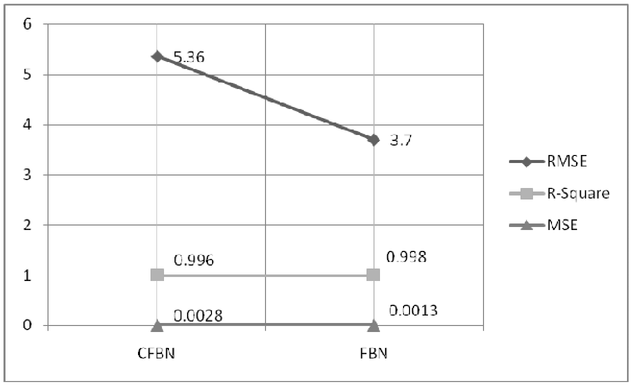



**Fig.3 Comparison of CF and FB artificial intelligence models**

For predicting sensory quality of instant coffee flavoured sterilized drink, the efficiency of different models was observed to be as follows:

FB and CF artificial intelligence models were used for mapping between inputs and outputs patterns. Various compositions of threshold functions were used in layers. Several topologies were tested and the best results which used from each network, training algorithm and threshold functions. The best results belonged to FBN model, LOGSIG-LOGSIG-LOGSIG as threshold function and Levenberg–Marquardt as learning algorithm. FBN model with single hidden layer having seventeen neurons gave the best results: (3.70% RMSE; 0.998 $R^2$ ; 0.0013 MSE) as represented in Table 1 and Fig. 2. For CFB artificial intelligence model, Levenberg–Marquardt was used as learning algorithm and TANSIG–LOGSIG –LOGSIG as threshold function. CF artificial intelligence model with single hidden layer having seven neurons: (5.36% RMSE; 0.996 $R^2$ ; 0.0028 MSE) gave best results as represented in Table 2 and Fig. 2. Thus, the study revealed that FB artificial intelligence model is better than CFB artificial intelligence model in prediction of sensory quality of instant coffee flavoured sterilized drink as represented in Fig. 3. The trend of our results is in agreement with the earlier findings of, Chayjan[10]and Phomkong [11] , who observed that for FB artificial intelligence model, LOGSIG-LOGSIG-LOGSIG as threshold function and Levenberg–Marquardt, and also for CF artificial intelligence model Levenberg–Marquardt used as learning algorithm and TANSIG– LOGSIG –LOGSIG as threshold function gave best results for modeling by soft computing.

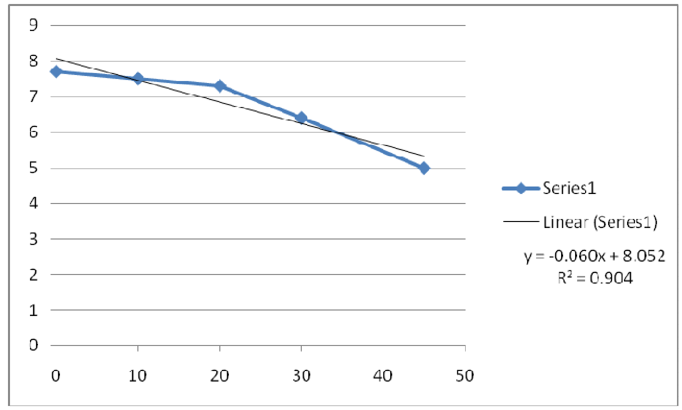



**Fig.4 Shelf life prediction using regression equations**

Regressions equations based on sensory scores and constant were developed. Regression coefficient came out as -0.6 and constant 8.0052, and $R^2$ was found to be 90 percent as represented in Fig.4, after solving them 0.551 came as the output which was subtracted from the experimental shelf life of instant coffee flavoured sterilized drink *i.e.*, 45 days and it came as 44.44 days.

## IV. CONCLUSION

Artificial neural network models based on feedforward backpropagation and cascade forward algorithms were developed for prediction of sensory quality of instant coffee flavoured sterilized drink. The performances of the two developed artificial intelligence models were compared with each other. Feedforward backpropagation artificial intelligence model exhibited best results (3.70% RMSE; 0.998 $R^2$; 0.0013 MSE), compared to cascade forward artificial intelligence model (5.36% RMSE; 0.996 $R^2$ ; 0.0028 MSE). From the study, it is concluded that the developed feedforward backpropagation model is better than cascade forward artificial intelligence model in prediction of sensory quality of instant coffee flavoured sterilized drink. The shelf life predicted 44.44 days is very close to the experimentally obtained shelf life of 45 days.

## REFERENCES

[1] A.Araque,E.D. Martin, G. Perea, J.I. Arellano and W. Buno. "Synaptically released acetylcholine evokes Ca2+ elevations in astrocytes in hippocampal slices". *Journal of Neuroscience*, vol. 22, pp. 2443-2450,2002.

[2] P.D. Wasserman. "Advanced Methods in Neural Computing", Van Nostrand Reinhold, New York, 255. 1993.

[3] R.S. Parmer, R.W. McClendon, G. Hoogenboom, P.D. Blankenship, R.J. Cole and J.W. Dorner." Estimation of aflatoxin contamination in preharvest peanuts using neural networks". *Transaction ASAE* vol.40, no.3,pp. 809–813,1997.

[4] S.M. Goñi, S. Oddone, J.A. Segura, R.H. Mascheroni and V.O. Salvadori

." Prediction of foods freezing and thawing times: Artificial neural networks and genetic algorithm approach". *Journal of Food Process Engineering*, vol.84, no.1, pp. 164-178, 2008.

[5] Sumit Goyal and G.K. Goyal ." Application of artificial neural engineering and regression models for forecasting shelf life of instant coffee drink". *International Journal of Computer Science Issues*. (accepted for publication Vol. 8, no. 4, 2011)

[6] Sumit Goyal and G.K. Goyal. " Development of Intelligent Computing Expert System Models for Shelf Life Prediction of Soft Mouth Melting Milk Cakes".*International Journal of Computer Applications*. (accepted for publication in July 2011 Issue).

[7] Sumit Goyal and G.K Goyal."Central Nervous System Based Intelligent Computer Engineering System for Shelf Life Detection of Soft Mouth Melting Milk Cakes" *Journal of Computing* (accepted for publication in Vol. 3, no. 7, 2011)

[8] http://en.wikipedia.org/wiki/Shelf_life (accessed on 28.6.2011).

[9] H.Demuth, M. Beale and M.Hagan. "Neural Network Toolbox User's Guide". The MathWorks, Inc., Natrick, USA. 2009

[10] R.A. Chayjan." Modeling of sesame seed dehydration energy requirements by a soft-computing". *Australian journal of crop science*,vol. 4, no.3, pp.180-184,2010.

[11] W. Phomkong, G. Srzedninchi and R.H. Driscoll. Desorption isotherms of stone fruit. *Drying Technology*, vol. 24, pp. 201-210, 2006.

**BIOGRAPHIES**

**First author's Biodata :** Sumit Goyal received Bachelor of Information Technology degree in 2005 and Masters in Computer Applications in 2007 from central university of Government of India, IGNOU, New Delhi. He received his Master of Philosophy degree in computer science in the year 2008 from Vinayak Missions University, Tamil Nadu, India. He has been working as Sr. Research Fellow at National Dairy Research Institute, Karnal, India. His research interests have been in the area of artificial neural networks and prediction of shelf life of food products.

**Second author's Biodata:**

Gyanendra Kumar Goyal received the B.S. Degree in 1970 and the PhD. degree in 1979 from Panjab University, Chandigarh, India. From 1975 to 2009, he was Scientist, Sr. Scientist and Principal Scientist at the National Dairy Research Institute, Karnal, India. During 1985 – 86, he did specialization in Dairy and Food Packaging from School of Packaging, Michigan State University, Michigan, East Lansing, USA; and during 1999 he specialized in Education Technology at Cornell University, New York, USA. Presently he is working as Emeritus Scientist at National Dairy Research Institute, Karnal, India. His research interest includes dairy & food packaging and shelf life determination.