



Study on Single and Double Hidden Layers of Cascade Artificial Neural Intelligence Neurocomputing Models for Predicting Sensory Quality of Roasted Coffee Flavoured Sterilized Drink

Sumit Goyal
Senior Research Fellow
National Dairy Research Institute,
Karnal-132001 (Haryana), India

Gyanendra Kumar Goyal
Emeritus Scientist
National Dairy Research Institute,
Karnal-132001 (Haryana), India

ABSTRACT

For centuries, coffee has been brewed and consumed in households, hot shops and restaurants. Today flavoured milks have become very popular and they contain nutrients as compared with soft drinks. Sterilized milk is the product made by heating milk to high temperature (121°C) with 15 m holding time so that it remains fit for human consumption for longer time at room temperature. Efficiency of single and double hidden layers of Cascade neurocomputing models for prediction of sensory quality of roasted coffee flavoured sterilized drink were studied. Colour and appearance, viscosity, flavour and sediment were taken as input parameters, while overall acceptability was used as output parameter. The results of cascade neurocomputing models were calculated with two types of prediction performance measures, viz., root mean square error and coefficient of determination R^2 . The study revealed that more the number of neurons in single hidden layer, less the error for cascade neurocomputing models (RMSE:0.00011; R^2 : 0.999999; neurons:50).

General Terms

Artificial Neural Network, MATLAB, Soft Computing

Keywords

ANN, Cascade, Neurocomputing, Sensory Quality, Coffee, Prediction

1. INTRODUCTION

Since long, coffee has been brewed and consumed in households, hot shops and restaurants. Coffee as beverage is prepared from the roasted seeds (beans) of the coffee plant. Today flavoured milks have become very popular as they contain nutrients compared to soft drinks, and are manufactured by using natural and synthetic flavours, viz., chocolate, banana, vanilla, strawberry, raspberry and orange. Sterilized milk is the product made by heating milk to high temperature (121°C) with 15 m holding time so that it remains fit for human consumption for longer period at room temperature. Neurocomputing model is composed of many single elements called neurons. Artificial Neural Networks (ANN) were born after McCulloch and Pitts introduced a set of simplified neurons in the year 1943 (Learnartificialneuralnetworks web-site, 2011). Neurons are basic signaling units of the nervous system of a living being in which each neuron is a discrete cell whose several processes are from its cell body. These neurons are connected to each

other in different ways, forming different types of neurocomputing models. The most popular ANN is the multilayer feed-forward neural network where the neurons are arranged into layers: input layer, hidden layer, and output layer (Bardot *et al.*, 1994) as represented in Fig. 1. Neurocomputing model usually has one or more hidden layers, which enable the network to deal with the non-linear and complex correlation (Masters, 1993).

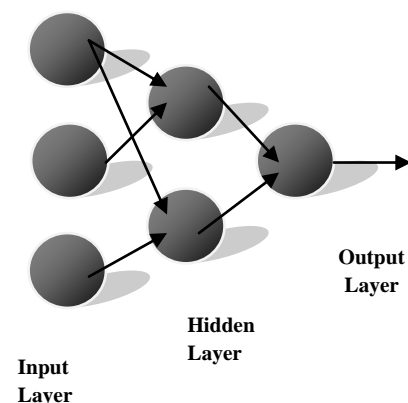


Fig.1. Single hidden layer model

1.1 The Input Layer

The input layer to the neural network is the conduit through which the external environment presents a pattern to the neural network. Once a pattern is presented to the input layer of the neural network, the output layer produces another pattern. The input layer represents the condition for which neural network is getting trained. Every input neuron represents some independent variable that has an influence over the output of the neural network (Heatonresearch web-site, 2011).

1.2 The output layer

The output layer of the neural network presents a pattern to the external environment. The number of output neurons directly relate to the type of work neural network performs. If the neural network is used for classifyinf items into groups, then it is better to have one output of neurons for each group. If the neural network is expected noise reduction on a signal, then the number of input neurons

match the number of output neurons (Heatonresearch web-site, 2011).

1.3 The number of hidden layers

The number of hidden layers to be taken in the neural network should be decided followed by determining the number of neurons in each layer (Heatonresearch web-site, 2011). Neural networks with two hidden layers can represent functions with any kind of shape as shown in Fig. 2.

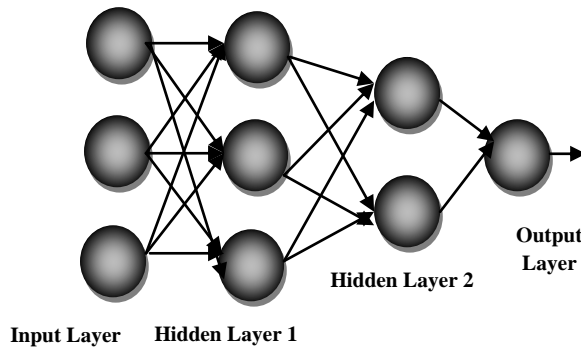


Fig.2. Network with 2 hidden layers

Table 1: Differences between the numbers of hidden layers

Number of Hidden Layers	Result
1	Can approximate arbitrarily while any functions which contains a continuous mapping from one finite space to another
2	Represent an arbitrary decision boundary to arbitrary accuracy with rational activation functions and can approximate any smooth mapping to any accuracy

Applying very few neurons in the hidden layers results in underfitting, as they are unable to adequately detect the signals in a complicated dataset, while use of numerous neurons in the hidden layers causes *overfitting*, and increase in training time of the network. For determining the optimum number of neurons in the hidden layers, following methods have been recommended (Tradingsolutions web-site, 2011):

- The number of hidden neurons should be in the range between the size of the input layer and the size of the output layer.
- The number of hidden neurons should be 2/3 of the input layer size, plus the size of the output layer.

- The number of hidden neurons should be less than twice the input layer size.

Thus the architecture of neural network comes down to trial and error approach for determining the numbers of hidden neurons, *i.e.*, “forward” and “backward” selection methods. In “forward selection method”, small number of hidden neurons is taken usually beginning with only two hidden neurons, and the neural network is trained and tested. The number of hidden neurons is then increased in systematic manner, and the process is repeated till the results of the training and testing improve. In “backward selection method”, after taking large number of hidden neurons, the neural network is trained and tested repeatedly till there is no further improvement in the performance of the neural network (Heatonresearch web-site, 2011). Goyal and Goyal (2011a) implemented artificial neural engineering and regressions models for forecasting shelf life of instant coffee drink. Goyal and Goyal (2011b) applied linear layer (design) and time - delay methods of intelligent computing expert system for shelf life prediction of soft mouth melting milk cakes. Neuron based Elman and Generalized Regression models for estimating shelf life of instant coffee sterilized drink were developed by Goyal and Goyal (2011c). They concluded that neuron based computing models are efficient in predicting shelf life of instant coffee sterilized drink. Goyal and Goyal (2011d) applied Elman and self-organizing simulated neural network models for predicting shelf life of soft cakes. The objective of the present work was to study efficiency of single and double hidden layers cascade forward (CF) neurocomputing models for predicting sensory quality of roasted coffee flavoured sterilized drink.

2. MATERIAL AND METHODS

2.1 Data modeling

The dataset consisted of 50 observations, randomly divided into two disjoint subsets, *i.e.*, training set having 40 observations (80% of total observations) and test set comprising of 10 observations (20% of total observations). Colour and appearance, flavour, viscosity and sediment were taken as input parameters. The Overall acceptability was used as output parameter for developing the neurocomputing models (Fig.3).

2.2 Training and validation

80% of the observations were used for training and 20% of the observations were used for validation of cascade and feedforward neurocomputing models. The dataset consisted of two parts, *i.e.*, input parameters (colour and appearance, flavour, viscosity and sediment), and output parameter (overall acceptability). Fig.4 represents the training pattern of the models used in developing neurocomputing models as shown in Fig.3.

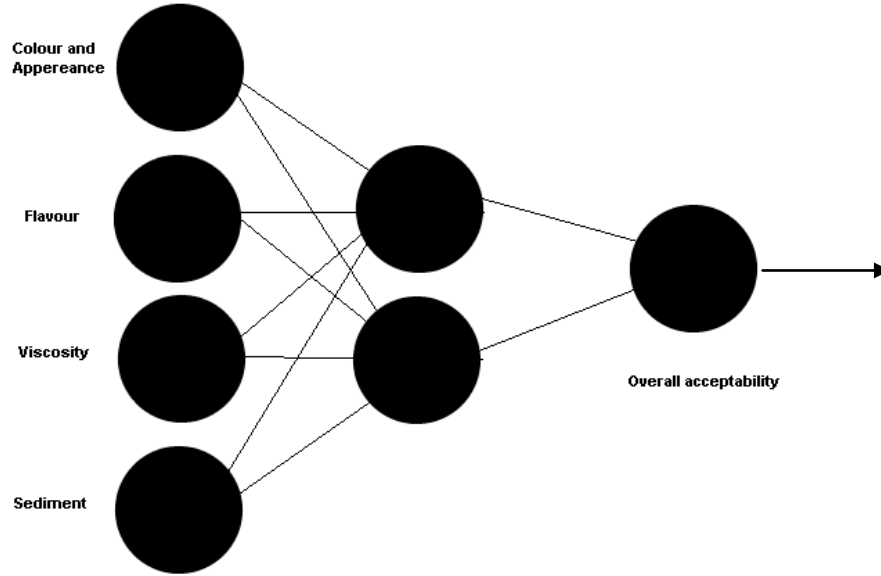


Fig.3. Input and output parameters for neurocomputing models

The data was normalized to the range of 0 and 1, then utilized to train cascade neurocomputing models by using MATLAB software. Single as well as double hidden layers were selected for evaluating the performance of neurocomputing models. Number of neurons in each hidden layer varied from 1 to 50. The sum square errors were performance function. Weights and biases were randomly initialized. The network was trained with 100 epochs and transfer function for each hidden layer was tangent sigmoid, while for the output layer, it was pure linear function. Tangent sigmoid transfer function maps the input to the interval (-1, 1). Each pair of input and target output was applied for training the weights and biases of neurocomputing models with tangent sigmoid function. Numerous backpropagation algorithms were investigated like Fletcher–Reeves update conjugate gradient algorithm, gradient descent algorithm with adaptive learning rate, Levenberg–Marquardt algorithm, Polak– Ribière Update conjugate gradient algorithm, Powell–Beale restarts conjugate gradient algorithm, BFG quasi-Newton algorithm, and Bayesian regularization. Bayesian regularization gave good results compared to other algorithms; therefore it was used as training function. RMSE (1) and R^2 (2) were used for comparing the prediction performance of both the neurocomputing models.

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

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

Where, Q_{exp} = Observed value; Q_{cal} = Predicted value; n = Number of observations in dataset.

Selection of weights



Training network



Calculating error



Weights adjustment



Minimum



End

Fig. 4. Training pattern of models

3. RESULTS AND DISCUSSION

Table 1. Cascade forward neurocomputing models with single hidden layer

Neurons in Hidden Layer	RMSE	R^2
4	0.0021	0.999953
9	0.0022	0.999947
13	0.0022	0.999949
20	0.0158	0.997382
35	0.0031	0.999899
40	0.0012	0.999983
50	0.0001	0.999999



Table 2. Cascade forward neurocomputing models with double hidden layer

Neurons in Hidden Layer	RMSE	R^2
5,4	0.003116611	0.999903078
5,5	0.002431899	0.999941176
7,4	0.001784197	0.999968304
8,7	0.005466342	0.999699464
15,15	0.002171112	0.999952619

Cascade neurocomputing models were developed for testing efficiency of single and double hidden layers for prediction of sensory quality of roasted coffee flavoured sterilized drink. Numerous topologies were tried and tested, as there is no defined method to reach to a good conclusion other than hit and trial approach. Several constituents of threshold functions were used in layers, but good results for feedforward neurocomputing model were obtained by combination of TANSIG-TRAINBR-PURELIN as threshold function and Bayesian regularization as learning algorithm. The experiments were carried out with single as well as double hidden layers for both the models (Table 1 and Table 2). Cascade neurocomputing models are similar to feedforward networks, but include a weight connection from the input to each layer and from each layer to the successive layers (Demuth *et al.*, 2009). For cascade neurocomputing model, Bayesian regularization was used as learning algorithm and TANSIG- TRAINBR –PURELIN as threshold function. Cascade neurocomputing model with single hidden layer having fifty neurons gave best results (RMSE 0.00011, R^2 0.999999) (Table 1), thus confirming the earlier reported observations (Demuth *et al.*, 2009) that more the number of neurons in single hidden layer, less is the error for cascade neurocomputing models.

4. CONCLUSION

Cascade neurocomputing models were developed for predicting sensory quality of roasted coffee flavoured sterilized drink. Different combinations were tried and tested. As the number of neurons increased in hidden layers, also the training time. From the study it is concluded that more number of neurons in single hidden layer results in less error for cascade neurocomputing models (RMSE: 0.00011; R^2 : 0.999999; neurons:50).

5. REFERENCES

- [1] Bardot. I, Bouchereau, L., Martin, N. and Alagos , B. 1994. Sensory instrumental correlation by combining data analysis and neural network techniques. Food Quality and Preference, 51(2), 159–166.
- [2] Demuth.H., Beale, M. and Hagan, M. 2009. Neural Network Toolbox User's Guide. The MathWorks, Inc., Natick, USA.
- [3] Goyal, Sumit and Goyal, G.K. 2011a. Application of artificial neural engineering and regression models for forecasting shelf life of instant coffee drink. International Journal of Computer Science Issues, 8(4), No 1, pp.320-324.
- [4] Goyal, Sumit and Goyal, G.K. 2011b. Development of intelligent computing expert system models for shelf life prediction of soft mouth melting milk cakes. International Journal of Computer Applications, 25(9), pp.41-44.
- [5] Goyal. Sumit and Goyal, G.K. 2011c. Development of neuron based artificial intelligent scientific computer engineering models for estimating shelf life of instant coffee sterilized drink. International Journal of Computational Intelligence and Information Security, 2(7), pp.4-12.
- [6] Goyal, Sumit and Goyal, G.K. 2011d. Simulated neural network intelligent computing models for predicting shelf life of soft cakes. Global Journal of Computer Science and Technology, 11(14), Version 1.0, pp. 29-33.
- [7] Learnartificialneuralnetworks web-site: <http://www.learnartificialneuralnetworks.com/> (accessed on 30.4.2011).
- [8] Masters, T. 1993. Practical neural network recipes in C++. London: Academic Press.
- [9] Heatonresearch web-site: <http://www.heatonresearch.com/articles/5/page2.html> (accessed on 19.8.2011).
- [10] Tradingsolutions website: http://www.tradingsolutions.com/webhelp/image/ebd_ebd452.gif . (accessed on 30.8.2011).