Optimized Neural Network Model to Characterize the Effects of Process Parameters on the Separation Efficiency of Iron Ore by a High Intensity Magnetic Separator

SK Das*, KK Bhattacharyya and Ratnakar Singh

CSIR-National Metallurgical Laboratory, Jamshedpur 831007, INDIA

Keywords: Neural Network; Optimized network architecture; Magnetic Separation; Recovery ; Grade; Feed Composition

Abstract

An improved and optimized multi-input-multi-output (MIMO) neural network model has been developed to predict the output parameters e.g grade and recovery to characterize the separation behavior of a high intensity magnetic separator for processing iron ore in the particle size range of 75~300 μm. The input parameters in the Neural model comprises of feed composition, % Fe, % SiO_2, %Al_2O_3 and process parameters such as particle size, pulp density and magnetic field intensity. The effect of process parameters on the separation efficiency was characterized by conducting a sensitivity analysis. The neural network architecture has been optimized using an efficient gradient based network optimization algorithm to minimize the training error rapidly. The model is based on the data generated from WHIMS experimental investigations. There has been an excellent agreement between the optimized model predictions with the measured values pertaining to recovery and grade for magnetic separation. This is depicted by the regression fit generated between the predicted and measured values.

.

*Corresponding author : Dr SK Das, E-Mail: skd@nmlindia.org , Ph. no: +91-657-2345085.

1. Introduction

The phenomena of magnetism and magnetic behaviour of materials have allowed the process of magnetic separation to be successfully employed in industrial processing. A wide variety of magnetic separation systems exist that has been used for industrial beneficiation processes for many years. Whether by a lifting, trapping or deflection technique, a magnetic field is generated that will selectively act upon one material in preference to another by virtue of their different magnetic response. Due to the paramagnetic properties of hematite ore, in that its magnetic susceptibility naturally increases in an increasingly powerful magnetic field, one technique utilized in fine iron ore recovery is wet high intensity magnetic separation or the WHIMS. Advances in High Gradient Magnetic Separation (HGMS) [1] in the early 1980's aided in the arena of fine paramagnetic particle removal. Quantitative determination of separation behavior in mineral processing operations in terms of process and operating parameters based on first principal modeling has remained inadequate due to a complex and non-linear relationship between the separation characteristics such as recovery, yield and grade (inputs) as a function of process and operating variables (output). Artificial neural network (ANN) is a learning technique that enables to map the hidden input-output relationship accurately. Although there have been few neural network models applied to various unit operations [2,3,4] in mineral processing, namely, hydrocyclone, flotation and grinding mills, however, ANN model based prediction for characterizing magnetic separation (WHIMS) of iron ore has been very scanty. The objective of the present work is to develop a multi-input-multi-output (MIMO) ANN model based on multilayer perceptron to predict recovery and grade of iron ore as a function of magnetic field intensity, % solid (pulp density), particle size and chemical composition of ore. In this proposed ANN model, a state-of-the-art efficient network optimization algorithm, namely, an improved Broyden-Fletcher-Goldfarb-Shanno (BFGS) have been incorporated to minimize the training error

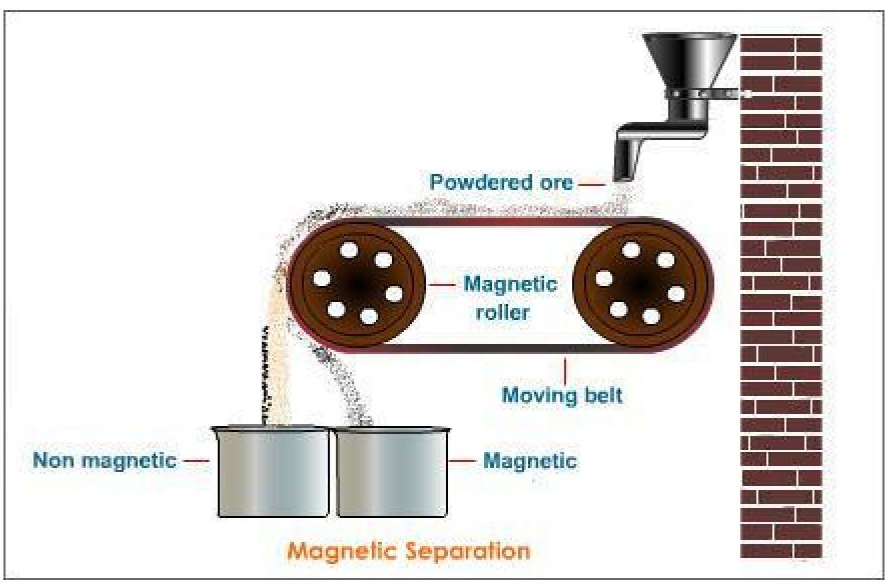

Fig. 1: Schematic of a High Intensity Magnetic Separator.

2. The building blocks of Neural Network Model

Artificial neural networks (ANNs) are non-linear self adaptive approach as opposed to traditional model based methods. They are powerful tools for modeling, specially when the underlying data relationship is unknown. ANN's can identify and learn correlated patterns between input data sets and corresponding target values. After training, ANN can be used to predict the outcome of new impendent input data. ANN imitate the learning process of human brain and can process problems involving non-liner and complex data even if the data are imprecise and noisy. The fundamental elements of the ANN methodology comprises of: (i) the functionality between input-output of neurons; (ii) the topological structure of the network; and (iii) the values of the connected weights and thresholds of neurons. MLP is an interconnection of perceptions in which data and calculations flow in a single direction, from the input data to the outputs. Figure 2 shows the conceptual framework of a MLP neural network architecture with input, hidden and output layers respectively.

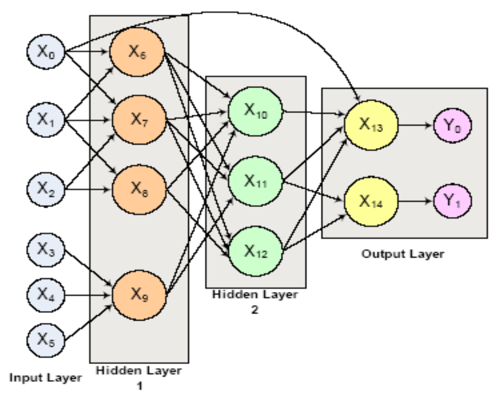

Fig 2: Typical Neural network architecture

The number of layers in a neural network is the number of layers of perceptions. The output from a given neuron is calculated by applying a transfer function to a weighted summation of its input to give an output, which can serve as input to other neurons. Mathematically this can be given as :

$$\alpha_{jk} = f_k\left(\sum_{i=1}^{N_{k-1}} w_{ijk}\alpha_{i(k-1)} + \beta_{jk}\right) \quad \ldots (1)$$

where, α_{jk} is neuron j's output from k's layer β_{jk} is the bias weight for neuron j in layer k. The model fitting parameters w_{ijk} are the connection weights and f_k's are activation functions.

2.2 Neural Network Training philosophy

The basic philosophy of training of neural networks seeks to minimize the error function

$$E = \sum_{p=1}^{N_p} \sum_{i=1}^{N_o} \left(t_{p,i} - A_{p,i}\right)^2 \qquad \text{........ (2)}$$

where N_p is the number of input/output data pairs, N_o is the number of outputs of the network, $t_{p,i}$ is the desired output value for a set of inputs and $A_{p,i}$ is the activation function for a given set of input**.** The partial derivative of the error function with respect to each weight is

$$\frac{\partial E}{\partial w_{j,i}} = 2 \sum_{p=1}^{N_p} \sum_{i=1}^{N_o} \left(t_{p,i} - A_{p,i}\right)\left(\frac{\partial t_{p,i}}{\partial w} - \frac{\partial A_{p,i}}{\partial w}\right) \qquad \text{...... (3)}$$

Weights are therefore adjusted proportionally to the negative of the error with respect to each weight.

$$\Delta w_{j \to i} = -\varepsilon \frac{\partial E}{\partial w_{j \to i}} \qquad \text{......... (4)}$$

where j is the sender node and i is the receiver node. ε is the learning rate of the network.

3. Gradient Based Network Optimization Method

Optimization of ANN's are concerned with the minimization of a particular objective function with respect to certain constraints. ANN's are proven highly efficient optimization tools. The objective of the network training is to find the optimal weights to minimize the errors between the prediction and the actual response. There are many different types of neural networks, differing by their network topology and or learning algorithm. Back Propagation (BP) learning and network optimization algorithm, which is one of the most commonly used algorithms is designed to predict the output parameters [5,6]. Training uses one of several possible optimization methods to minimize this error term. There are various back propagation algorithms such as Scaled Conjugate Gradient (SCG), [7] Levenberg-Marquardt (LM), Gradient Descent with Momentum (GDM), variable learning rate Back propagation (GDA) and Resilient back Propagation (RP). There is variety of network optimization techniques that uses gradient of a function to be optimized. The most recently developed highly efficient version of the quasi-Newton optimization methods is the BFGS algorithm, which has largely replaced the classical Davidson-Fletcher- Powell (DFP) algorithms. In general, the quasi-Newton method was derived from quadratic objective function. The inverse of the Hessian matrix, H is used to bias the gradient direction. $B = H^{-1}$ **...** (5)

In the quasi-Newton training method, the weights are updated using the following iterative procedure,

$$W_{i+1} = W_i - \eta B_i g_i \qquad \textbf{...} \text{ (6)}$$

The matrix B here need not be computed. It is successively estimated employing rank 1 or rank 2 updates, following each line search in a sequence of search directions. This is algorithmically given as follows

$$B_i = B_i - \Delta B_i \qquad \text{... (7)}$$

In this iterative algorithm, B_{i-1} is the previous value of B.

The ***BFGS algorithm*** can be invoked as [8] :

$$\Delta B_i = \left(1 + \frac{\Delta g^T B_i - 1\Delta g}{d^T \Delta g}\right)\frac{dd^T}{d^T \Delta g} - \frac{d\Delta g^T \quad B_i - 1 + B_i - 1\Delta g d^T}{d^T \Delta g} \qquad \text{... (8)}$$

Where, $d = w_i - w_i - 1$ and $\Delta_g = g_i - g_i - 1$, $\Delta B = B_i - B_{i-1}$ ....(9)

The BFGS algorithm has the advantage over DFP in that it does not require accurate line minimizations along the quasi-Newton directions to build up the approximate Hessian matrix. Thus, BFGS potentially reduces the number of function evaluations required to achieve an optimization

4. Results and Discussion

The optimized neural model input variables are namely, magnetic field intensity, % solid (pulp density), particle size and composition of ore (% Fe, % SiO_2, % Al_2O_3). Correspondingly, the output variables are namely, recovery (%) and grade (% Fe) of the ore. In the present neural model, input data set is segmented into three subsets, namely, one for training (learning),one for selection (validation), and one for testing (prediction) using roughly 2:1:1 ratio. Out of 300 datasets, 200 datasets are used as training samples, 50 as validation samples, and the remaining 50 samples have been utilized for prediction. In order to obtain the optimum network, 20 networks are first designed with six input neurons, two output neurons, and two different hidden layers (11 and 6) is considered for the network.

In the simulation results, fig 3 and 4 depict neural network predictions for grade and recovery, of iron ore using a WHIMS magnetic separation process. In these figures neural prediction are based on two best multi layer network architectures (MLP 6-11-2 and MLP 6-6-2) and are compared with the regression fit between predicted and measured outputs for the specified iron ore. Figures 5 and 6 show the predicted recovery and grade as a function of measured magnetic field intensity respectively during separation operation using the optimized architecture. It may be observed that the grade and recovery increase in a linearly as the magnetic field intensity is enhanced. Figures 7 and 8 describe the predicted behavior of grade and recovery as a function of % solid (pulp density) in the feed. The grade and recovery are found to increase with the higher concentration solid in the feed which is consistent with the operational observations. Figure 9 illustrates the neural prediction of recovery as a function of both % Fe and SiO_2 content of the ore. The surface topography of the 3-D plot indicates that the recovery is enhanced with higher Fe and lower silica content.

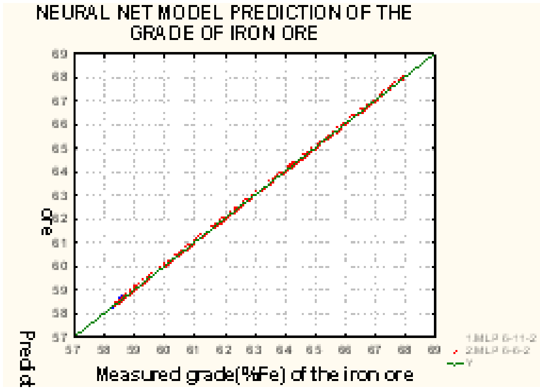

Fig. 3: Neural prediction and validation of the grade iron ore during magnetic separation (optimized network).

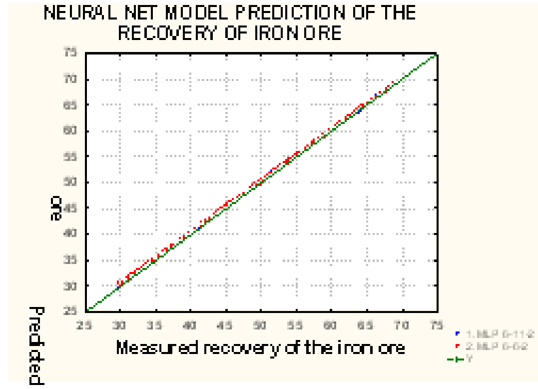

Fig. 4 Neural prediction and validation of the iron ore recovery during magnetic separation (optimized network).

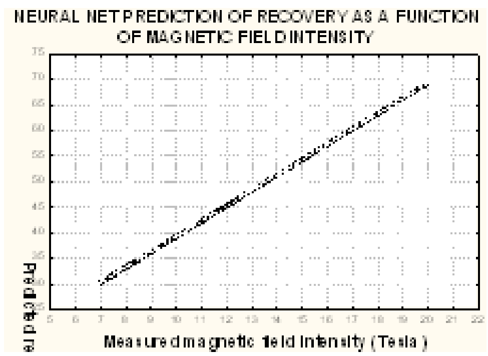

Fig. 5 Neural prediction of the recovery of iron ore as a function of measured magnetic field intensity (optimized network).

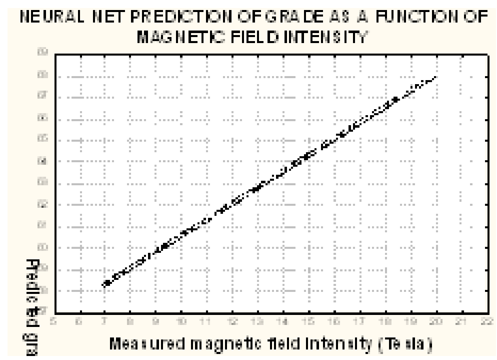

Fig. 6: Neural prediction of the grade of iron ore as a function of measured magnetic field intensity (optimized network).

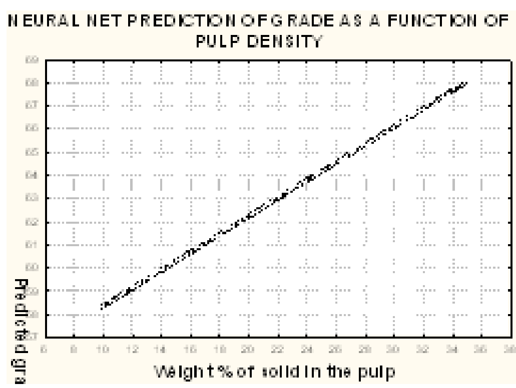

Fig. 7: Neural prediction of the grade of iron ore as a function of measured pulp density (optimized network)

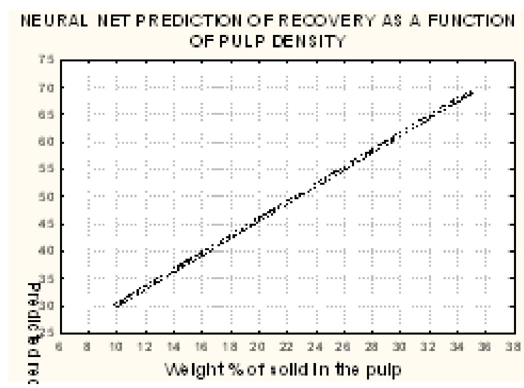

Fig. 8: Neural prediction of the recovery of iron ore as a function on measured pulp density (optimized network).

5. Conclusion

This paper provides a brief ANN modeling framework for prediction of separation characteristics of iron ore by magnetic separation (WHIMS) process with optimized network training Results of the neural network computer simulations are compared with those obtained from measured data. The main conclusions are as follows. The proposed optimized neural network model with a larger data set provides a sufficiently accurate predictive framework and compare extremely well with the measured data. The neural network approach provides an alternative modeling paradigm and has the advantage over other first principle based models treating polluted data or the data with strong non-linear relationships

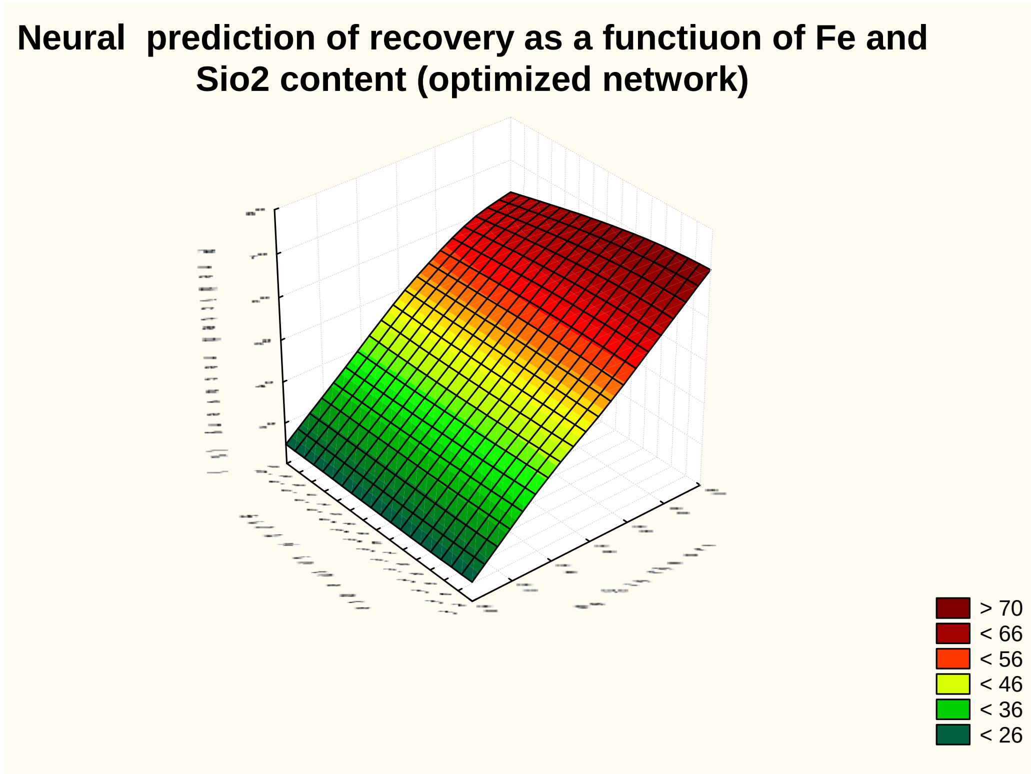

Fig 9: Neural prediction of recovery as a function of silica and Fe content of the ore (optimized network)

5. References

[1] Oberteuffer, J. A., 1973, *IEEE Trans. Magnetics*, MAG-**9**, pp. 303.

[2] Stange, W., 1993, *Minerals Engineering*, **6**(5), pp. 479.

[3] Van der Walt T. J., Van Deventer J. S. J. and Barnard E., 1993, *Minerals Engineering*, **6**(11), pp. 1135.

[4] Topalov A.V., Kaynak O., 2004, *Journal of Process Control*, **14**, pp. 581.

[5] Rumelhart D. E., Richard D., Richard G., Chauvin Y., 1995, Back Propagation: Theory, Architectures, and Applications, First Addition, Chapter 1, Lawrence Erlbaum Associates. Hillsdale, NJ. Publication (USA), p. 1.

[6] Trana C., Abrahamb A., Jaina L., 2004, *Neurocomputing* , **61**, pp. 85.

[7] Nash S. G., Nocedal J., 1991, *SIAM Journal of Optimization*, **1**, pp. 358.

[8] YUAN Y. X., 1991, *IMA Journal of Numerical Analysis*, **11**(3), pp. 325.