Neural Network Sensitivity Analysis Applied for the Reduction of the Sensor Matrix

Przemyslaw M. Szecówka[1], Andrzej Szczurek[2], Maciej A. Mazurowski[1], Benedykt W. Licznerski[1], and Franz Pichler[3]

[1] Faculty of Microsystem Electronics and Photonics, Wroclaw University of Technology, Janiszewskiego 11/17, 50-372 Wroclaw, Poland
przemyslaw.szecowka@pwr.wroc.pl
[2] Institute of Environmental Protection Engineering, Wroclaw University of Technology, Wybrzeze Wyspianskiego 27, 50-370 Wroclaw, Poland
[3] Institut fur Systemwissenschaften, Johannes Kepler University, Altenbergerstr. 69, A-4040 Linz, Austria

Abstract. The neural network sensitivity analysis, involving neural network training and the calculation of its outputs derivative on inputs, was applied to select the least significant sensor in the multicomponent gas mixtures analysis system. The sensitivity analysis results, collected for various neural network structures were compared with the real significances of the sensors, determined experimentally. The question of the influence of the correlation of the input vector elements on the analysis results was also illustrated and discussed.

1 Introduction

Contemporary gas mixtures analysis technology relies on the matrices of sensing elements and smart data processing techniques applied for their responses analysis, providing the desired information of qualitative or quantitative character. This approach is usually forced by the low selectivity of sensors, which disables simple calibration of 1 sensor for 1 gas appearing in the mixture. It may be observed that most of gas sensor systems described in the literature, although very successful, contain more or less redundant sets of sensing elements [1,2,3,4,5]. Shall be noted that each redundant sensor, applied in the matrix increases the cost of both fabrication and operation of the prospective system. Seems like the main problem is the lack of the reasonable and efficient methods of sensors selection.

If to assume that the preliminary version of the system, providing the acceptable accuracy of measurements is available (what is btw. usually reached using the large enough sensor array) the problem may be transformed to the elimination of the most redundant sensors, as far as the required performance of the system is preserved. The possible solution may be neural network sensitivity analysis [6,7], adopted for the estimation of the significance of the information given to the system by the particular sensors in the matrix. The neural networks

R. Moreno Díaz et al. (Eds.): EUROCAST 2005, LNCS 3643, pp. 27–32, 2005.

approach is somewhat unusual here. In the initial phase of the system construction the dummy neural network is trained to provide the sensitivity analysis and judge each sensor. After selection of the reasonable set of sensors the eventual data processing algorithm may be created using either neural network again or any other methodology.

The paper focused on the investigation of the efficiency and reliability of the neural network sensitivity analysis approach in the context of both the real world application in gas sensor system and the data artificially created for the purposes of the experiment.

2 Neural Network Sensitivity Analysis

The operation of the feedforward neural network with a single hidden layer and the sigmoid transfer function applied may be described by (1).

$$y_k^{(2)} = f\left(\sum_{j=0}^{J} w_{kj}^{(2)} f\left(\sum_{i=0}^{I} w_{ji}^{(1)} u_i\right)\right), \text{ where } f(x) = \frac{1}{1+e^{-x}} \tag{1}$$

After the training process, when the weights, denoted by w (with the appropriate indexes) are fixed, the neural network gains the unique approximation capabilities [8]. In the context of sensor systems the vectors of sensors responses u are transformed to the series of outputs $y^{(2)}$, providing desired information of either qualitative or quantitative character, on request.

The k-th neural network output sensitivity for the selected input u_i is defined as a derivative (2), which for the presumed construction (1) gives (3). Calculation of the sensitivity is made for each output-input pair and for each input pattern $u^{(p)}$. Concerning the patterns, the global sensitivity for the whole data set is calculated using for instance the Euclidean formula (4) or by finding the maximum absolute value. Eventually the sensitivity matrix is obtained with e.g. inputs listed in columns and outputs listed in rows. Further analysis may involve the min-max procedure providing the series of parameters describing how much the neural network is sensitive to the particular input. The inputs (i.e. the sensors in our context), with the lower values of sensitivity shall be considered as the candidates to remove.

$$s_{ki} = \frac{dy_k}{du_i} \tag{2}$$

$$s_{ki} = f'\left(x_k^{(2)}\right) \sum_{j=1}^{J} \left(w_{kj}^{(2)} f'\left(x_j^{(1)}\right) w_{ji}^{(1)}\right) \tag{3}$$

$$S_{ki} = \sqrt{\sum_{p=1}^{P} \left(s_{ki}^{(p)}\right)^2} \tag{4}$$

Formula (3) was originally proposed for pruning the redundant inputs of the neural network with the single hidden layer only [6]. The experience with construction of neural networks for the gas sensor arrays shows the need of applying the structures with two hidden layers to obtain the highest performance. Some doubts may appear then whether the sensitivity analysis applied for the too scant neural network structures would give the reliable results. Eventually the extended sensitivity formula, for the neural networks with two hidden layers, was calculated. Starting from (5) describing the appropriate neural network with two hidden layers, where u_h and $y_k^{(3)}$ denote the selected input and output, and the $w^{(1)}$, $w^{(2)}$ and $w^{(3)}$ with the appropriate indexes are the weights of the neurons in the following layers, the eventual sensitivity is given by (6).

$$y_k^{(3)} = f\left(\sum_{j=0}^{J} w_{kj}^{(3)} f\left(\sum_{i=0}^{I} w_{ji}^{(2)} f\left(\sum_{h=0}^{H} u_h\right)\right)\right) \tag{5}$$

$$s_{kh} = \frac{dy_k^{(3)}}{du_h} = f'\left(x_k^{(3)}\right) \sum_{j=1}^{J} \left(w_{kj}^{(3)} f'\left(x_j^{(2)}\right) \sum_{i=1}^{I} \left(w_{ji}^{(2)} f'\left(x_i^{(1)}\right) w_{ih}^{(1)}\right)\right) \tag{6}$$

Analysis of (3) and (6) induces that the extension of the sensitivity formula for the bigger and bigger neural networks may be generalized to the recurrence, somewhat similar to the classic error backpropagation [9], where the sensitivity of the bigger structure may be calculated as a weighted sum of the sensitivities calculated for the appropriate nodes of the smaller structure (7),(8).

$$s_{ki}^{(L)} = \frac{\partial y_k^{(L)}}{\partial u_i} = f'\left(x_k^{(L)}\right) \sum_{j=1}^{J} \left(s_{ji}^{(L-1)} w_{kj}^{(L)}\right) \tag{7}$$

$$s_{ki}^{(1)} = \frac{\partial y_k^{(1)}}{\partial u_i} = f'\left(x_k^{(1)}\right) w_{ki}^{(1)} \tag{8}$$

3 Experimental

The sensitivity analysis methodology was applied to the design of the sensor system providing quantitative analysis of the mixtures of butanol and toluene. The matrix, initially containing six sensors - TGS 800, TGS 822, TGS824, TGS 825, TGS 880, TGS 883 [10], was placed in a test chamber with controlled atmosphere for the multi-component characterisation. The set of vectors containing the gas concentrations and sensors responses was collected this way building up the 148 samples data set [2]. A series of the neural networks was created to provide some estimation of the reliablity of the method, with the sensors responses acting as the input and two gas concentrations as the desired output. The structures were varying between 6-10-2 and 6-40-2. For each neural network the sensitivity analysis was performed, i.e. sensitivities of all outputs to all the

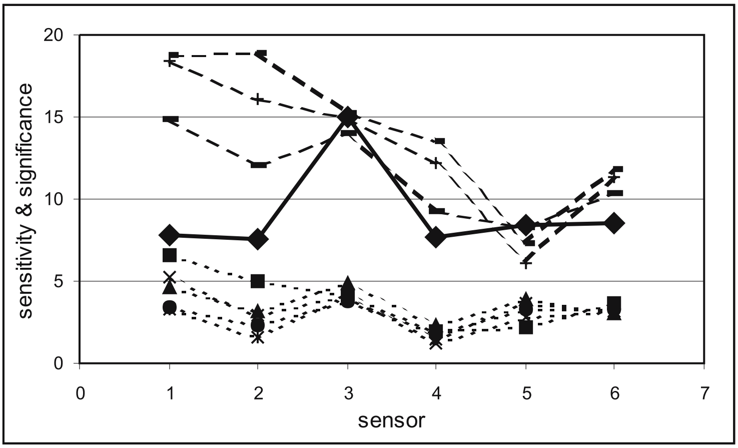

Fig. 1. Neural network sensitivities obtained for the structures with the single hidden layer (dotted lines) and with two hidden layers (dashed lines) compared with the experimental estimation of the sensors significance (thick solid line).

inputs were calculated for all the available samples. The global sensitivities for all patterns were calculated in two variants - Euclidean formula (5) and *maximum*. In-house developed software tools were used for both the neural networks development and sensitivity calculation. In further steps of the analysis the sensor with the lowest absolute value of the sensitivity factor shall be removed and the whole process may be repeated, with the reduced sensor matrix, to point the next one to remove etc. This process shall be stopped when either the system performance decreases dramatically or the results of the sensitivity analysis performed for series of neural networks are no longer coherent.

The sensitivity analysis performed for several neural networks with the single hidden layer consequently pointed to the sensor No. 4 as redundant. The details may be found in [11]. The sensitivity factors, obtained in several trials, are presented in Fig. 1 (the dotted lines). The dashed lines present analogous results obtained for the structures with two hidden layers. Various structures were implemented, starting from 6-12-8-2 up to 6-50-30-2. The results within this group are similar again, but this time sensor No. 5 is commonly recognized as redundant one. Such contradiction could be perceived as a stop condition for the sensor matrix reduction, but it is known from the other experiments that this set of sensors may be reduced indeed without visible loss in accuracy.

Further investigation of this phenomenon involved the introduction of the six data sets deriving from the original one, but with 1 input-sensor removed in each. The series of neural networks were trained for each variant, targeting in the experimental determination of the $significance$ factors for all the sensors. These factors were calculated as an average error of the 3 best structures. (The higher error of the neural network trained without particular input denotes its higher significance). The balance of these factors was plotted again in Fig. 1 (thick solid

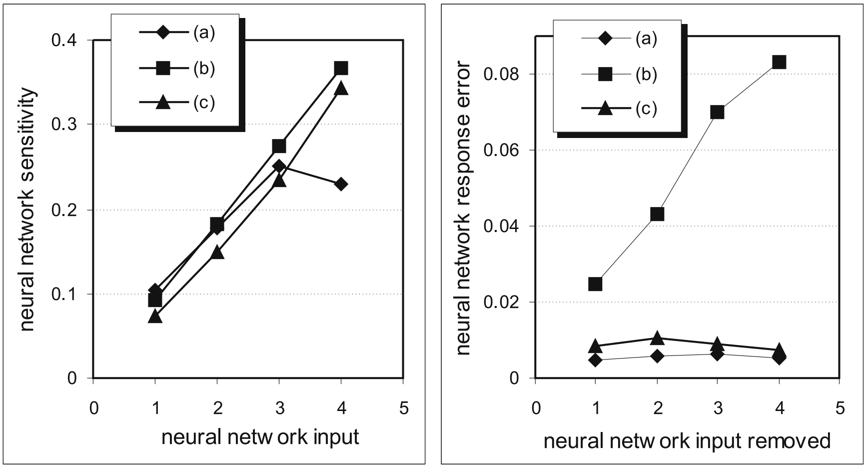

Fig. 2. Sensitivity analysis results (left) compared with the experimental estimation of the input significance (right) for (a) linear dependent, (b) independent and (c) non-linear dependent data

line). If to analyse the precise values of the sensors significance factors obtained this way, No. 2 shall be considered most redundant this time. Eventually the results of the sensitivity analysis performed for various neural networks, especially in the critical *first to remove* context are sometimes contradicting themselves and simultaneously contradicting the experimentally determined significance of the sensors. Shall be noted however that the experimental analysis, which shall be perceived as the most reliable here, estimates the significance of five sensors (i.e. No. 1, 2, 4, 5 and 6) at the very similar level. The insignificant differences mean that in fact any of these sensors could be removed, with similar impact on the system performance, what probably justifies the contradictions mentioned before.

The meaningless differences between the significance of the sensors and consequently the contradictions are probably caused by the correlation of the sensors responses (i.e the elements of the neural network input vector), which is very high. The simple experiment may show how the dependent inputs may keep the sensitivity analysis results far away from the real balance of the inputs significance. Let's take a sample function of $y = x_1 + 2x_2 + 3x_3 + 4x_4$ and generate a data set for the appropriate neural network training, in 3 variants - the first one with independent input variables x_1, x_2, x_3, x_4, the second one with linear dependence $x_4 = x_1 + x_2 + x_3$, and the third one with non-linear $x_4 = x_1^2 + x_2^2 + x_3^2$. The results of the sensitivity analysis are shown in Fig. 2 (on the left). These ones are very similar for all the data sets. And the real significance factors of the input variables, estimated by the "remove the input and train the neural network" procedure, are completely different for the dependent and non-dependent variants as it is shown in Fig. 2 (on the right), matching the intuitive expectations.

4 Conclusions

The neural network sensitivity analysis may be attractive tool for the reduction of the redundant sensor arrays. Presented experiments have shown however, that it does not provide the absolutely reliable results, when some elements of the input vector are dependent. It may be used, with care, as a reasonable heuristics for the construction of the effective gas sensor arrays, where the number of sensors is critical issue, and in many other fields requiring an estimation of the significance of the particular factor.

References

1. Gutierrez F.J., Ares L., Robla J. et al.: Integrated Sensors for Monitoring Contaminant Gases in Atmospheres and Soils, Proc. IEEE International Symposium on Industrial Electronics, Guimaraes, Portugal (1997) SS113- SS115.
2. Szczurek A., Szecówka P.M., Licznerski B.W.: Application of sensor array and neural networks for quantification of organic solvents vapours in air, Sensors and Actuators B, 58 (1999) 427-432.
3. Martin M.A., Santos J.P., Agapito J.A.: Application of artificial neural networks to calculate the partial gas concentrations in a mixture, Sensors and Actuators B 77 (2001) 681-471.
4. Sobański T., Saffarpour N.: Analysis of gas mixture using sensor array and artificial neural networks. In: Cybernetics and systems 2002. Proceedings of the Sixteenth European Meeting on Cybernetics and Systems Research. Vol. 1. (2002) 97-100.
5. Jervis B.W., Desfieux J., Jimenez J. et al.: Quantification of gas concentrations in mixtures of known gases using an array of different tin-oxide sensors, IEE P-Sci. Meas. Tech. 150 (3) (2003) 97-106.
6. Żurada J.M., Malinowski A., Usui S.: Perturbation method for deleting redundant inputs of perceptron networks, Neurocomputing 14 (1997) 177-193.
7. Engelbrecht A.P.: A New Pruning Heuristic Based on Variance Analysis of Sensitivity Information, IEEE Trans. On Neural Networks, Vol. 12, No. 6 (2001) 1386-1399.
8. Kornik K., Stinchcombe M., White H.: Multilayer feedforward networks are universal approximators, Neural Networks, vol. 2 (1989) 359-366.
9. Rumelhart D.E., Hinton G.E., Williams R.J.: Learning Representations by Back-Propagating Errors, Nature, 322, (1986) 533-536.
10. Figaro Gas Sensors, TGS 800 series specification, Figaro Engineering Inc.
11. Szecówka P.M., Szczurek A.: Sensors selection for gas mixtures analysis systems, Proc. 26-th International Spring Seminar on Electronic Technology, Stara Lesna (2003) 398-403.