ELM-Based Ensemble Classifier for Gas Sensor Array Drift Dataset

D. Arul Pon Daniel, K. Thangavel, R. Manavalan and R. Subash Chandra Boss

Abstract Much work has been done on classification for the past fifteen years to develop adapted techniques and robust algorithms. The problem of data correction in the presence of simultaneous sources of drift, other than sensor drift, should also be investigated, since it is often the case in practical situations. ELM is a competitive machine learning technique, which has been applied in different domains for classification. In this paper, ELM with different activation functions has been implemented for gas sensor array drift dataset. The experimental results show that the ELM with bipolar function classifies the drift dataset with an average accuracy of 96 % than the other function. The proposed method is compared with SVM.

Keywords ELM · Ensembles · Gas sensor array drift dataset · Bipolar

1 Introduction

The past decade has seen a significant increase in the application of multi-sensor arrays to gas classification and quantification. The idea to combine an array of sensors with a pattern recognition algorithm to improve the selectivity of the single

D. Arul Pon Daniel (✉) · K. Thangavel · R. Subash Chandra Boss
Department of Computer Science, Periyar University, Salem 636011, India
e-mail: apdaniel86@yahoo.com

K. Thangavel
e-mail: drktvelu@yahoo.com

R. Subash Chandra Boss
e-mail: rmsubash_18@yahoo.co.in

R. Manavalan
Department of Computer Application, KSR Arts and Science College, Trichengodu, India
e-mail: manavalan_r@rediffmail.com

G. S. S. Krishnan et al. (eds.), *Computational Intelligence, Cyber Security and Computational Models*, Advances in Intelligent Systems and Computing 246,
DOI: 10.1007/978-81-322-1680-3_10,

gas sensor has been widely accepted and being used by researchers in this field. In fact, an array of different gas sensors is used to generate a unique signature for each gas [1]. A single sensor in the array should not be highly specific in its response but should respond to a broad range of compounds, so that different patterns are expected to be related to different odors [2]. Different methods have been suggested recently to compensate for sensor drift in experiments for gas identification [3]. Chemical sensor arrays combined with read-out electronics and a properly trained pattern recognition stage are considered to be the candidate instrument to detect and recognize odors as gas mixtures and volatiles [4].

After learning the features of the class, the SVM recognizes unknown samples as a member of a specific class. SVMs have been shown to perform especially well in multiple areas of biological analyses, especially functional class prediction from microarray sensors produced data [5].

It is not surprising to see that it may take several minutes, several hours, and several days to train neural networks in most of the applications. Unlike traditional popular implementations, for single-hidden-layer feedforward neural networks (SLFNs) with additive neurons, which is a new learning algorithm called extreme learning machine (ELM) [6].

This paper has been organized into five sections. Section 2 presents the short note about the dataset used. Sect. 3 describes the approach of (ELM). In Sect. 4, experimental results of various activation function and discussion are presented. In Sect. 5, conclusions and further research scope are presented.

2 Dataset

The drift dataset contains 13,910 measurements from 16 chemical sensors utilized in simulations for drift compensation in a discrimination task of six gases at various levels of concentrations. The resulting dataset comprises of six distinct pure gaseous substances, namely ammonia, acetaldehyde, acetone, ethylene, ethanol, and toluene, each dosed at a wide variety of concentration values ranging from 5 to 1,000 ppmv [7]. This dataset is available in http://archive.ics.uci.edu/ml/datasets/Gas+Sensor+Array+Drift+Dataset

3 Extreme Learning Machine

Recently, a new learning algorithm for SLFN named the ELM has been proposed by Huang et al. [6, 8]. The SLFN with randomly chosen input weights and hidden bias can approximate any continuous function to any desirable accuracy. ELM is a

single-hidden-layer neural network with good generalization capabilities and extreme learning capacity. The generalization performance of ELM for classification problem depends on three parameters. Number of hidden nodes, the input weights, and the bias values are needed to be optimally chosen [8]. For hidden layer, many activation functions such as sigmoidal, sine, Gaussian, and hard limiting function can be used, and the output neurons have a linear function as an activation function [9–11].

The ELM has several interesting and significant features different from traditional popular learning algorithms.

For the dataset which contain N distinguish objects (x_i, t_i) where $x_i = [x_{i1}, x_{i2}, x_{i3}, \ldots, x_{in}]^T \in R^n$ and $t_i = [t_{i1}, t_{i2}, t_{i3}, \ldots, t_{im}]^T \in R^m$, the relationship between the actual output of SLFN, with an infinite differentiable activation function $g(x)$, and the target outputs t_i is given by

$$\sum_{i=1}^{\tilde{N}} \beta_i g(w_i x_i + b_i) = t_j, \quad j = 1, \ldots, N. \tag{1}$$

Here, $\tilde{N}$ is the number of hidden nodes, $w_i = [w_{i1}, w_{i2}, w_{i3}, \ldots, w_{in}]^T$ and $\beta_i = [\beta_{i1}, \beta_{i2}, \beta_{i3}, \ldots, \beta_{im}]^T$are the weight vector connecting inputs to the ith hidden neuron and the ith hidden neuron to output neurons, respectively, and b_i is the bias of the ith hidden neuron. Equation (1) can be rewritten compactly as $H\beta = T$

where $H = \begin{bmatrix} g(\mathrm{w}_1.\mathrm{x}_1 + \mathrm{b}_1) & \ldots & g(\mathrm{w}_{\tilde{\mathrm{N}}}.\mathrm{x}_1 + \mathrm{b}_{\tilde{\mathrm{N}}}) \\ . & \ldots & . \\ . & \ldots & . \\ . & \ldots & . \\ g(\mathrm{w}_i.\mathrm{x}_\mathrm{N} + b_1) & \ldots & g(\mathrm{w}_{\tilde{\mathrm{N}}}.\mathrm{x}_\mathrm{N} + \mathrm{b}_{\tilde{\mathrm{N}}}) \end{bmatrix}_{\mathrm{N}\times\tilde{\mathrm{N}}}$ is called the hidden layer output matrix of the neural network [6], and $\beta_i = [\beta_i, \ldots, \beta_{\tilde{N}}]^T_{\tilde{N}\times m}$, $T = [t_i, \ldots, t_N]^T_{N\times m}$.

Traditionally, training of SLFN has typically applied the back-propagation learning algorithm to adjust the set of weights (w_i, β_i) and biases (b_i). It is common and problem dependent.

ELM is introduced to resolve the issues in back-propagation neural network. Initially, parameters for the hidden node may be randomly specified. The output weights can then be analytically determined. Also, it is shown that the upper bound of the required number of hidden nodes is the number of distinct training objects (i.e., $\widetilde{N} \leq N$). Thus, given (pre-specified) N, associated with parameters (w_i, b_i), the hidden nodes can be randomly generated. Determining the output weights β is as simple as finding the least-square solutions to the given linear system.

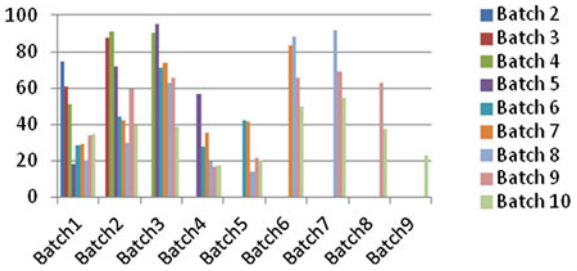

Fig. 1 Classification accuracy of the SVM classifiers with RBF kernel function

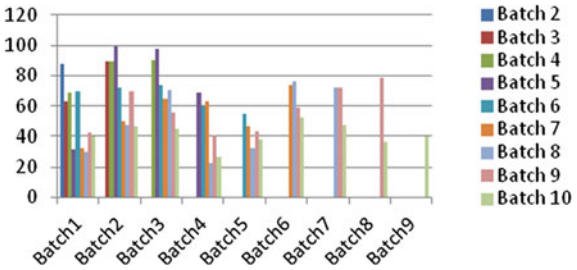

Fig. 2 Classification accuracy of the ELM classifiers with unipolar function

4 Experimental Results

In this experiment, the features in the training datasets are scaled appropriately to lie between −1 and +1. The kernel bandwidth parameter, the SVM parameter, and ELM parameter were chosen using 10-fold cross-validation by performing a grid search in the range $[2^{-10}, 2^{-9},\ldots, 2^{4}, 2^{5}]$ and $[2^{-5}, 2^{-4},\ldots, 2^{9}, 2^{10}]$, respectively. The performance of an SVM trained on batch 1 and tested on batches 2–10 respectively. Note that this curve is estimated with the same SVM model used in Fig. 1 but tested on data from batches instead of months. Similar behaviors were found when we trained several SVMs on batches 2–5 and tested them on successive batches. These results are again shown in Fig. 2. The complete set of results, i.e., the accuracy of classifiers trained on batches 1–9 and tested on successive batches, is given in Table 1. The individual plots correspond to the performance of classifier trained with batch 1 and tested on batches at subsequent time points after applying the component correction method for every one of the six reference gases (Figs. 3, 4 and 5).

In this section, the gas sensor array drift data are classified by the activation function unipolar, bipolar, and radial basis kernel, and they are classified into six classes. Before classifying, features are normalized between −1 and +1.

Table 1 Classification accuracy (in %) on batches 2–10 by SVM and ELM with activation function (unipolar, bipolar, RBF)

Batch ID	Function	Batch 2	Batch 3	Batch 4	Batch 5	Batch 6	Batch 7	Batch 8	Batch 9	Batch 10
Batch 1	SVM RBF	74.36	61.03	50.93	18.27	28.26	28.81	20.07	34.26	34.48
	ELM Unipolar	87.86	62.79	68.94	31.47	69.65	32.07	29.59	42.12	39.72
	ELM Bipolar	85.28	63.36	76.39	44.16	69.91	33.76	31.63	44.04	37.27
	ELM RBF	87.78	69.92	63.97	33.50	65.95	37.83	40.47	42.55	40.72
Batch 2	SVM RBF		87.83	90.68	72.08	44.52	42.46	29.93	59.57	39.69
	ELM Unipolar		89.28	89.44	98.98	72.08	49.51	47.27	69.36	46.13
	ELM Bipolar		84.23	77.64	80.71	73.52	46.00	61.90	60.42	40.61
	ELM RBF		73.70	73.91	89.84	70.43	45.64	59.86	60.85	43.72
Batch 3	SVM RBF			90.06	94.92	70.96	73.73	62.59	65.74	38.89
	ELM Unipolar			90.06	97.46	73.95	64.76	70.74	55.74	45.13
	ELM Bipolar			85.71	98.98	75.26	65.70	70.06	53.61	45.50
	ELM RBF			85.71	97.46	71.60	64.57	66.32	55.74	45.13
Batch 4	SVM RBF				56.35	27.52	35.40	19.73	17.02	17.56
	ELM Unipolar				69.03	60.04	63.35	22.44	40.85	26.41
	ELM Bipolar				98.47	63.91	62.77	32.99	49.36	44.30
	ELM RBF				92.89	63.3	68.53	26.87	42.55	42.86
Batch 5	SVM RBF					42.52	41.32	13.95	21.49	20.11
	ELM Unipolar					54.34	46.38	32.65	42.97	38.61
	ELM Bipolar					58.56	54.30	70.40	50.42	39.05
	ELM RBF					54.47	50.34	54.42	48.08	32.33
Batch 6	SVM RBF						83.53	88.44	65.74	49.97
	ELM Unipolar						73.42	75.85	58.51	52.13
	ELM Bipolar						72.82	77.55	82.12	53.52
	ELM RBF						70.13	74.83	62.76	51.47

(continued)

Table 1 (continued)

Batch ID	Function	Batch 2	Batch 3	Batch 4	Batch 5	Batch 6	Batch 7	Batch 8	Batch 9	Batch 10
Batch 7	SVM RBF							91.84	69.15	54.28
	ELM Unipolar							72.10	71.70	47.36
	ELM Bipolar							78.57	74.89	60.75
	ELM RBF							72.10	71.70	49.47
Batch 8	SVM RBF								62.98	37.69
	ELM Unipolar								78.93	36.86
	ELM Bipolar								83.83	42.00
	ELM RBF								81.27	40.08
Batch 9	SVM RBF									22.64
	ELM Unipolar									39.86
	ELM Bipolar									53.55
	ELM RBF									43.97

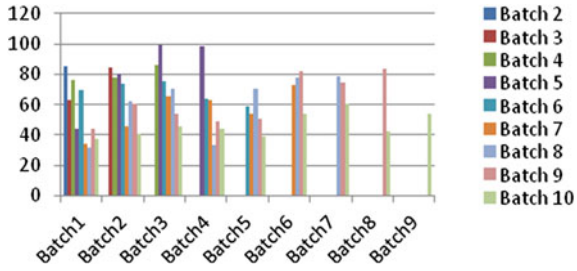

Fig. 3 Classification accuracy of the ELM classifiers with bipolar function

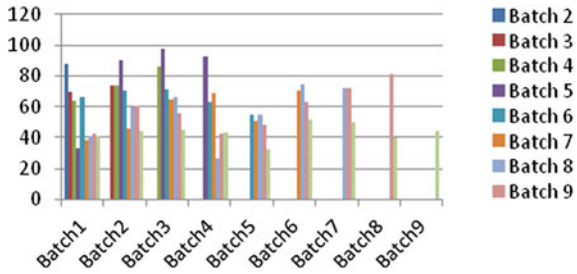

Fig. 4 Classification accuracy of the ELM classifiers with RBF kernel function

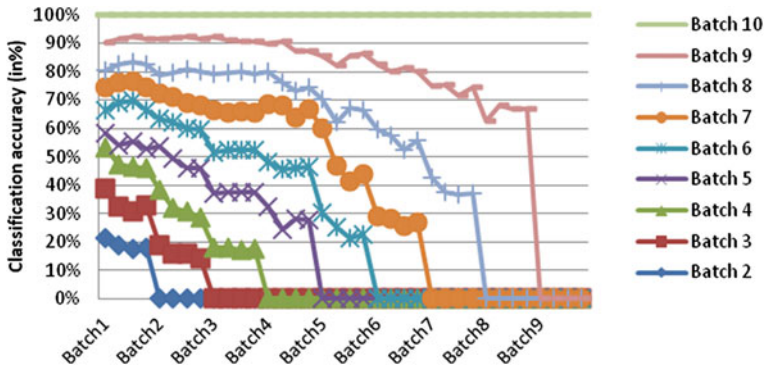

Fig. 5 Classification accuracy (in %) on batches 2–10 by SVM and ELM with activation function (unipolar, bipolar, RBF)

5 Conclusion

Gas sensor array drift dataset has been analyzed using SVM and the proposed ELM methods. Six chemical components are used to acquire the drift dataset with different time series. In this paper, ELM has been used for classification and compared with SVM. The proposed ELM method achieves the average accuracy of 92.23 % when compared with SVM. This classification of chemical components may be used to train the system to defect the cancer from human exhaled breathe in future.

Acknowledgments The first and fourth author immensely acknowledges the partial financial assistance under University Research Fellowship, Periyar University, Salem.

The second author immensely acknowledges the UGC, New Delhi, for partial financial assistance under UGC-SAP (DRS) Grant No. F.3-50/2011.

References

1. Sofiane Brahim-Belhouari, Amine Bermak and Philip C. H. Chan (2004) Gas Identification with Microelectronic Gas Sensor in Presence of Drift Using Robust GMM. IEEE ICASSP 2004, 0-7803-8484-9/04/$20.00, pp. V-833–V-836.
2. Arul Pon Daniel D, Thangavel K, and Subash Chandra Boss R (2012) A Review of Early De-tection of Cancers using Breath Analysis. Proc. IEEE Conf. Pattern Recognition, Infor-matics and Mobile Engineering (PRIME 2012), IEEE Press, DOI: 10.1109/ICPRIME.2013.6208385: 433–438.
3. John-Erik Haugen, Oliver Tomic, Knut Kvaal (1999) A calibration method for handling the temporal drift of solid state gas-sensors. Analytica Chimica Acta, pp. 23–39.
4. Persaud K, Dodd G (1982) Analysis of discrimination mechanisms in the mammalian olfactory system using a model nose. Nature 299 (5881):352–355.
5. Belusov AI, Verkazov SA, von Frese J (2002) Applicational aspects of support vector machines. J Chemometric, 16(8–12):482–489.
6. Huang G B, Zhu Q Y, Siew C K (2004) Extreme learning machine: A new learning scheme of feedforward neural networks. In Proceedings of International Joint Conference on Neural Networks, Budapest, Hungary, 2:985–990.
7. Alexander Vergara, Shankar Vembu, Tuba Ayhan, Margaret A. Ryan, Margie L. Homer and Ramón Huerta (2012) Chemical gas sensor drift compensation using classifier ensembles. Sensors and Actuators B: Chemical, DOI: 10.1016/j.snb.2012.01.074: 320–329.
8. Huang G, Zhu Q, Siew C (2006) Exreme Learning Machine: Theory and applications. Neuro-computing. 70(1-3):489–501.
9. Brady M, Highnam R (1999) Mammographic Image Analysis. Kluwer series on medical image Understanding.
10. Hassanien (2007) Fuzzy rough sets hybrid scheme for breast cancer detection. Image and Vision Computing, 25(2):172–183.
11. Roffilli M (2006) Advanced machine learning techniques for digital mammography. Technical Report UBLCS, University of Bologna. Italy.