

Design of an Artificial Neural Network Pattern Recognition Scheme Using Full Factorial Experiment

Ibrahim Masood^{1, a}, Nadia Zulikha Zainal Abidin^{1, b}, Nur Rashida Roshidi^{1, c},
Noor Azlina binti Rejab^{1, d}, Mohd Faizal Johari^{1, e}

¹ Faculty of Mechanical and Manufacturing Engineering, Universiti Tun Hussein Onn Malaysia,
86400 Parit Raja, Batu Pahat, Johor, Malaysia

^a ibrahim@uthm.edu.my , ^b nadiazulikha@gmail.com , ^c cdashieda@gmail.com ,
^d azlinarejab88@gmail.com , ^e faizalj@ gmail.com

Keywords: Artificial neural network, full factorial design of experiment, multi-model classifier, multivariate quality control, pattern recognition

Abstract. Automated recognition of process variation patterns using an artificial neural network (ANN) model classifier is a useful technique for multivariate quality control. Proper design of the classifier is critical for achieving effective recognition performance (RP). The existing classifiers were mainly designed empirically. In this research, full factorial design of experiment was utilized for investigating the effect of four design parameters, i.e., recognition window size, training data amount, training data quality and hidden neuron amount. The pattern recognition study focuses on bivariate correlated process mean shifts for cross correlation function, $\rho = 0.1 \sim 0.9$ and mean shifts, $\mu = \pm 0.75 \sim 3.00$ standard deviations. Raw data was used as input representation for a generalized model ANN classifier. The findings suggested that: (i) the best performance for each pattern could be achieved by setting different design parameters through specific classifiers, which (ii) gave superior result (average RP = 98.85%) compared to an empirical design (average RP = 96.5%). This research has provided a new perspective in designing ANN pattern recognition scheme in the field of statistical process control.

Introduction

In manufacturing industries, process variation has become a major source of poor quality. When manufacturing process involves two or more correlated variables (multivariate), an appropriate scheme is necessary to monitor these variables jointly. In recent years, this issue has been addressed through the development of artificial neural networks (ANN) based pattern recognition schemes. Various schemes such as Novelty Detector ANN [1], Ensemble-ANN [2], Multi-Module-Structure-ANN [3] and Features-based ANN [4] have been proposed for automatically recognizing multivariate process variation patterns. Further discussion on these schemes can be found in reference [5, 6].

In general, the ANN models were utilized as pattern classifier. As such, a proper design of the classifier is critical in achieving efficient recognition performance (RP). The above literatures reported that the classifier performance was influenced by several design parameters such as network architecture, patterns behaviour, training data amount and training algorithm. For case by case basis, they were commonly determined empirically (based on experience, self assumption, or trial and error experiments). In this study, full factorial design of experiment (FFDOE) was utilized for investigating an optimal design parameters for a generalized ANN model classifier. Details discussion is organized as follows: framework of an ANN scheme; results and discussion on experimental design; conclusions.

ANN Pattern Recognition Scheme

Framework for an ANN pattern recognition scheme was developed based on single stage monitoring and diagnosis approach as shown in Figure 1. Process monitoring refers to the identification of process status either it is running within a statistically in-control or out-of-control

condition, whereas process diagnosis refers to the identification of the source variables of out-of-control process in mean shifts.

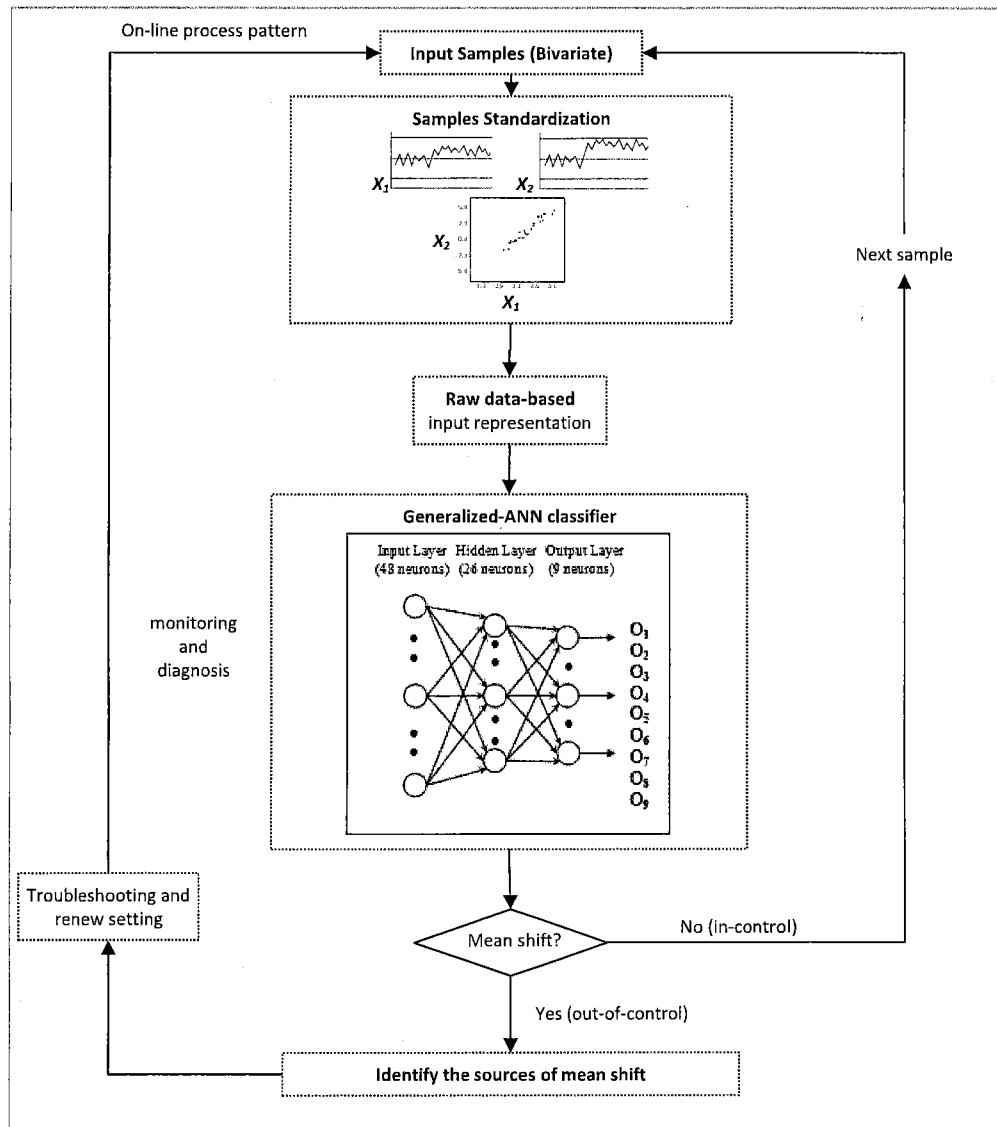


Figure 1. Framework of an ANN pattern recognition scheme

Modeling of data of multivariate process variation patterns, design, training and testing of the generalized ANN classifier can be referred in reference [4]. Process patterns were classified as normal and abnormal patterns (upward and downward shifts). Ideally, raw data should be tapped from real manufacturing industry. However, since a large number of data are required, the artificial data were generated using Lehman model [7]. This methodology has been adopted in previous research [1, 4].

“Partially developed shift patterns” and “dynamic patterns” were applied into the ANN training and testing respectively to suit for on-line (dynamic) process condition. Tables 1 and 2 provide the details parameters for modeling artificial data of partially developed shift patterns. In order to achieve the best training result for overall pattern categories, the amount of training patterns were set as follows: (i) bivariate normal patterns = $[1500 \times (\text{total combination of data correlation})]$ and (ii) bivariate shift patterns = $[100 \times (\text{total combination of mean shifts}) \times (\text{total$

combinations of data correlation)](refer to Table 1). In order to improve discrimination capability between normal and shift patterns, a huge amount of N(0,0) patterns was applied into ANN training. The US(1,1) and DS(1,1) pattern categories also require a huge amount of training patterns since it contain a more complex combination of mean shifts compared to the other bivariate shift pattern categories. In order to achieve the best monitoring performance (not discussed in this article), different percentage of partially developed shift patterns were utilized for different range of mean shifts as shown in Table 2. The starting points of sudden shifts (SS) were determined empirically. The actual value of data correlation is dependent to the variability in the bivariate data. The simulated values ($\rho = 0.1, 0.3, 0.5, 0.7, 0.9$) could only be achieved when the process data streams are in either “fully normal pattern” or in “fully developed shift pattern”.

For representing into an ANN classifier, the samples were standardized to a range between $[-3, +3]$ and then normalized to a compact range between $[-1, +1]$. For normalization, a maximum and a minimum values were taken from the overall standardized samples. The normalization could minimize noise from the samples, thus provides accurate and consistent recognition.

Table 1. Parameters for the artificial data

Pattern category	Mean shift (in std dev)	Data correlation (ρ)	Amount of training patterns
N(0, 0)	X1: 0.00 X2: 0.00	0.1, 0.3, 0.5, 0.7, 0.9	1,500 x 5 = 7,500
US(1, 0)	X1: 1.00, 1.25, ..., 3.00 X2: 0.00, 0.00, ..., 0.00		100 x 9 x 5 = 4,500
US(0, 1)	X2: 0.00, 0.00, ..., 0.00 X1: 1.00, 1.25, ..., 3.00		100 x 9 x 5 = 4,500
US(1, 1)	X1: 1.00, 1.00, 1.25, 1.25, ..., 3.00 X2: 1.00, 1.25, 1.00, 1.25, ..., 3.00		100 x 25 x 5 = 12,500
DS(1, 0)	X1: -1.00, -1.25, ..., -3.00 X2: 0.00, 0.00, ..., 0.00		100 x 9 x 5 = 4,500
DS(0, 1)	X2: 0.00, 0.00, ..., 0.00 X1: -1.00, -1.25, ..., -3.00		100 x 9 x 5 = 4,500
DS(1, 1)	X1: -1.00, -1.00, -1.25, -1.25, ..., -3.00 X2: -1.00, -1.25, -1.00, -1.25, ..., -3.00		100 x 25 x 5 = 12,500
USDS	X1: 1.00, 1.00, 1.25, 1.25, ..., 3.00 X2: -1.00, -1.25, -1.00, -1.25, ..., -3.00		100 x 25 x 5 = 12,500
DSUS	X1: -1.00, -1.00, -1.25, -1.25, ..., -3.00 X2: 1.00, 1.25, 1.00, 1.25, ..., 3.00		100 x 25 x 5 = 12,500

Table 2. Parameters for the partially developed shift patterns

Range of mean shifts (in standard deviations)	Amount of partially developed shift patterns	Starting point of sudden shift (SS)
$\pm [1.00, 1.50]$	2/3 (66.7%)	Sample 9 th
$\pm [1.75, 2.25]$	1/2 (50.0 %)	Sample 13 th
$\pm [2.50, 3.00]$	1/3 (33.3 %)	Sample 17 th

ANN classifier was trained using gradient descent with momentum and adaptive learning rate (traingdx). Maximum number of epochs and error goal were set as 500 and 0.001 respectively. Learning rate and learning rate increment were set as 0.05 and 1.05, whereas maximum number of epochs and error goal were set as 1500 and 0.001 respectively. Network performance was evaluated based on mean square error (MSE). The hyperbolic tangent function in hidden layer limited the

hidden output between $[-1, +1]$ and sigmoid function in output layer limited the classification output between $[0, 1]$. The training process was stopped whenever either the error goal has been met or the maximum allowable training epoch has been reached. The training reached error goal between 100 to 300 epochs.

Experimental Design: Results and Discussion

FFDOE was utilized in investigating the effects of four design parameters as shown in Table 3 towards the RP of a generalized ANN model. Then, an optimal design was proposed. The analysis was performed using MINITAB software which involved selection of experiment design and main effect plot. Table 4 shows the overall results of experiments.

Table 3. Design Parameters for an ANN model

Design Parameters	Relationship to ANN	Low Level	High Level
Recognition window size (WS)	Size for input layer	12	26
Number of hidden neuron (HD)	Size for hidden layer	10	30
Amount of training dataset (DS)	Recognition capability	50 (per shifts magnitude)	150 (per shifts magnitude)
Quality of abnormal patterns (QD)	Recognition capability	0.5 std dev (shifts interval)	0.25 std dev (shifts interval)

Table 4. FFDOE selection and results

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14	C15	C16	C17
	StdOrder	RunOrder	CenterPt	Blocks	WS	HD	DS	QD	W00	US10	US01	US11	DS10	DS01	DS11	USDS	DSUS
1	13	1	1	1	12	10	150	9	99.53	90.54	85.30	98.63	85.94	89.13	98.09	99.53	99.68
2	9	2	1	1	12	10	50	9	98.33	82.96	95.80	99.15	90.68	91.46	99.40	99.43	98.68
3	15	3	1	1	12	30	150	9	95.84	58.65	53.81	94.67	39.36	48.18	95.15	96.76	96.67
4	6	4	1	1	26	10	150	5	97.51	89.06	95.90	96.04	83.18	79.77	96.32	96.56	98.18
5	2	5	1	1	26	10	50	5	97.51	96.82	94.44	98.23	97.88	94.19	98.70	99.25	98.48
6	4	6	1	1	26	30	50	5	77.36	75.28	83.09	99.57	80.58	81.00	97.29	98.88	99.89
7	10	7	1	1	26	10	50	9	98.45	74.77	73.60	96.01	82.17	68.57	95.91	96.99	96.42
8	11	8	1	1	12	30	50	9	99.13	92.37	98.76	99.26	97.39	97.65	99.37	99.84	99.22
9	5	9	1	1	12	10	150	5	97.24	86.33	83.59	95.84	82.94	81.52	96.94	97.69	96.45
10	14	10	1	1	26	10	150	9	99.56	94.87	94.24	99.50	97.00	94.86	99.11	99.53	99.67
11	16	11	1	1	26	30	150	9	98.45	80.49	82.55	97.04	82.47	85.33	95.92	96.99	97.37
12	8	12	1	1	26	30	150	5	91.22	97.22	96.74	98.58	98.04	96.38	99.54	99.68	99.29
13	7	13	1	1	12	30	150	5	99.73	41.86	45.78	89.09	39.03	46.07	86.94	95.05	94.51
14	12	14	1	1	26	30	50	9	95.77	48.48	70.20	98.94	53.51	73.20	98.11	99.73	98.96
15	3	15	1	1	12	30	50	5	98.15	56.54	74.07	95.53	69.17	70.71	95.40	97.35	98.16
16	1	16	1	1	12	10	50	5	99.25	83.78	91.33	99.00	80.80	94.16	97.85	99.03	99.17
17																	

Based on main effects plot analysis as shown in Figure 2, WS can be observed as the most effective parameter, among others, for overall shift patterns being recognized. This parameter

performed better performance at high level (26). Parameter HD at low level (10) gave better performance for overall shift patterns. Increment in HD impaired the recognition accuracy (e.g. US10 patterns: HD = 10, RP = 97.22%; HD = 30, RP = 94.87%). Meanwhile, in recognizing normal pattern (N00), this parameter indicated better result at high level (30). As such, in order to balance the discrimination capability between normal and abnormal patterns, an exact value of HD could be within a range of [10, 30]. Parameter QD, on the other hand, gave an inverse effect in comparison with HD. This parameter performed better for overall shifts patterns at high level (0.25 standard deviations of shifts interval). Inversely, it gave better result at low level (0.50 standard deviations of shifts interval) for normal patterns. Parameter DS has different effects for different patterns being recognized. This parameter performed better at low level (50 per shifts magnitude) for N00, US01, US10, DS10 and DS01. Inversely, it performed better at high level for US11, DS11, USDS and DSUS patterns.

Table 5 summarizes the guideline of design parameter setting for achieving efficient recognition accuracy performance (RP) for each individual pattern. For instance, the highest performance for normal pattern (N00) could be achieved by setting the parameters at WS = 26, HD = 30, DS = 50 and QD = 5. This guideline represents the outcomes of FFDOE analysis. Based on this guideline, further experiments have been performed to evaluate the performance of an ANN model classifier when design with FFDOE compared to an empirical design. Obviously, the ANN model with FFDOE has outperformed the empirical designed ANN. Since the different type of patterns could be recognized effectively using specific design parameters setting, the best performance could be realized through the development of multi-model ANN classifiers (specific classifier). The proposed multi-model should consist of 9 individual ANN classifiers which are trained using specific design parameter setting.

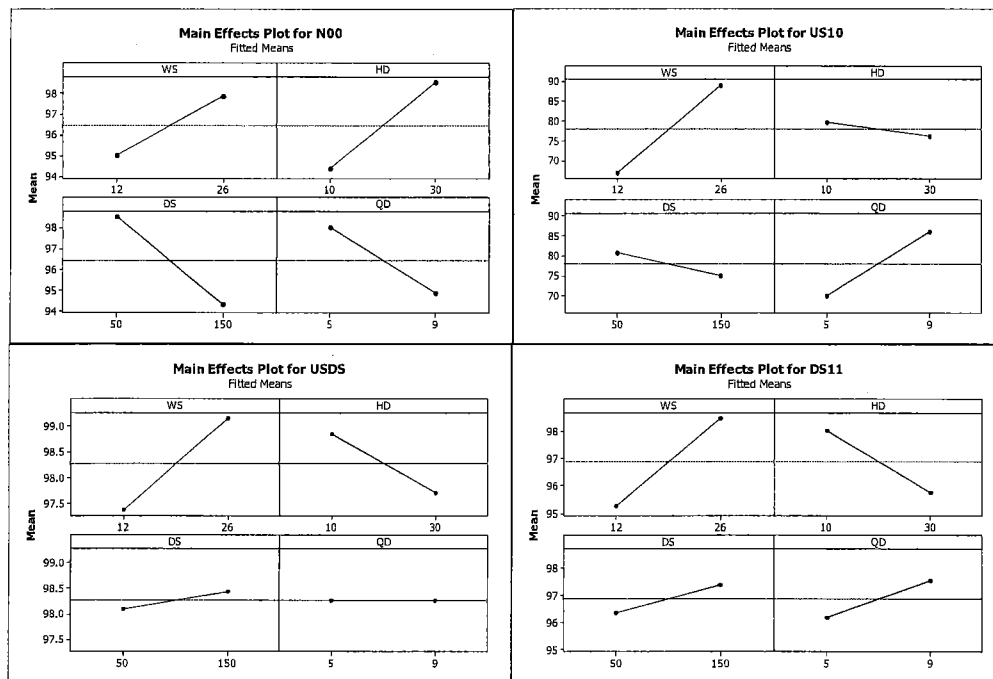


Figure 2. Main effects plot for several individual patterns

Table 5. Design parameter setting for recognizing individual patterns

Pattern	FFDOE								Empirical Design	
	WS		HD		DS		QD		RP (%)	RP (%)
	12	26	10	30	50	150	5	9		
N00		✓		✓	✓		✓		99.73	97.72
US10		✓	✓		✓			✓	97.17	90.6
US01		✓	✓		✓			✓	98.76	93.35
US11		✓	✓			✓		✓	99.52	99.25
DS10		✓	✓		✓			✓	97.89	94.54
DS01		✓	✓		✓			✓	97.65	96.04
DS11		✓	✓			✓		✓	99.54	98.76
USDS		✓	✓			✓		✓	99.68	99.25
DSUS		✓	✓			✓		✓	99.70	99.03

Note: Empirical design - WS = 24, HD = 26, DS = 100, QD = 9

Conclusions

In developing a pattern recognition scheme for multivariate quality control application, proper selection of design parameters of an ANN model classifier is important towards heightening the recognition performance. In previous researches, the design parameters were mainly determined empirically. This study, however, proved that the classical FFDOE with main effect plot analysis was effective in designing an ANN classifier. The findings suggested that: (i) Different types of patterns such as normal and shifts could be recognized effectively using specific design parameters. As such, the best performance for each pattern could be achieved through multi-model ANN classifier, and (ii) FFDOE gave superior results in comparison with an empirical design. Overall, a guideline in designing ANN pattern recognition scheme could be a reference in the field of statistical process control.

References

- [1] Zorriassatine F, Tannock JDT, O'Brien C (2003) Using novelty detection to identify abnormalities caused by mean shifts in bivariate processes. *Computers and Industrial Engineering* 44: 385–408.
- [2] Yu JB, Xi LF, Zhou XJ (2009) Identifying source(s) of out-of-control signals in multivariate manufacturing processes using selective neural network ensemble. *Engineering Applications of Artificial Intelligence* 22: 141–152.
- [3] El-Midany TT, El-Baz MA, Abd-Elwahed MS (2010) A proposed framework for control chart pattern recognition in multivariate process using artificial neural networks. *Expert Systems with Applications* 37: 1035–1042.
- [4] Masood I, Hassan A (2013) Pattern recognition for bivariate process mean shifts using feature-based artificial neural network. *International Journal of Advanced Manufacturing Technology* 66 (9–12): 1201–1218.
- [5] Masood I, Hassan A (2010) Issues in development of artificial neural network-based control chart pattern recognition schemes. *European Journal of Scientific Research* 39(3): 336–355.
- [6] Hachica W, Ghorbel A (2012) A survey of control-chart pattern-recognition literature (1991–2010) based on a new conceptual classification scheme. *Computers and Industrial Engineering* 63: 204–222.
- [7] Lehman RS (1977) *Computer simulation and modeling: an introduction*. Lawrence Erlbaum, London