Comparing Statistical Feature and Artificial Neural Networks for Control Chart Pattern Recognition: A Case Study

M. Arbi Hadiyat [1], and Kestrilia Rega Prilianti. [2]

[1] Industrial Engineering, Universitas Surabaya

[2] Bioinformatics Interest, Technobiology Faculty, Universitas Surabaya

Abstract. Control chart has been widely used for monitoring production process, especially in evaluating the quality performance of a product. An uncontrolled process is usually known by recognizing its chart pattern, and then performing some actions to overcome the problems. In high speed production process, real-time data is recorded and plotted almost automatically, and the control chart pattern needs to be recognized immediately for detecting any unusual process behavior. Neural networks for automatic control chart recognition have been studied in detecting its pattern. In the field of computer science, the performance of its automatic and fast recognition ability can be a substitution for a conventional method by human. Some researchers even have developed newer algorithm to increase the recognition process of this neural networks control chart. However, artificial approaches have some difficulties in implementation, especially due to its sophisticated programming algorithm. Another competing method, based on statistical feature also has been considered in recognition process. Control chart is related to applied statistical method, so it is not unreasonable if statistical properties are developed for its pattern recognition. Correlation coefficient, one of classic statistical features, can be applied in control chart recognition. It is a simpler approach than the artificial one. In this paper, the comparison between these two methods starts by evaluating the behavior of control chart time series point, and measured for its closeness to some training data that are generated by simulation and followed some unusual control chart pattern. For both methods, the performance is evaluated by comparing their ability in detecting the pattern of generated control chart points. As a sophisticated method, neural networks give better recognition ability. The statistical features method simply calculate the correlation coefficient, even with small differences in recognizing the generated pattern compared to neural networks, but provides easy interpretation to justify the unusual control chart pattern. Both methods are then applied in a case study and performances are then measured.

Keywords: Control chart, pattern recognition, neural network, correlation, back propagation

1. Introduction

One of main tools in Statistical Process Control (SPC) is control chart, which gives information of production process behaviour. Mean shift, unusual, cyclical and increasing mean patterns of the control chart could be detected as informative feedback for later improvement. Manual detection for unusual process behaviour integrated with some statistical package (i.e. Minitab or SPSS) can be handled for normal production speed, with the real-time data taken from process can be manually inputted to. Of course, for the high speed production, this semi-manual real-data input could be particular problem, especially when there are numerous data to be plotted in control chart. Automatic real-data retrieve control chart has been widely used to capture production process behaviour pattern, as developed firstly by Wiel et al. [8] that was applying visual pattern monitoring manner. As the technology increases, recent methods for recognizing the control chart pattern have been developed and replaced the manually visual monitoring way.

Artificial intelligence for pattern recognition has been being a key in the recognition process and producing many multi-discipline researches on it. Anagun [2] and Ghanim [1] have applied standard feed forward neural networks for control chart pattern recognition and showed its ability in pattern detection. Some researchers even combine neural networks optimization algorithm with some features to increase the recognition process, such as statistics Hassan et al. [5], and genetic algorithm Ebrahimzadeh and Ranaee [3]. Further, the learning process in neural networks also improved by featuring the data inputted to, see Masood and Hassan [6]. All those researches deal with the advanced method requiring complicated programming ability.

On the other side, the practical applications of control chart only need better recognition process without considering the methods used inside. Complicated method couldn't assure the best result in interpreting the control chart behaviour, as a simple method couldn't too. Correlation coefficient, a simple statistical feature firstly applied by Yang and Yang [9] for control chart pattern recognition. As a simple one, this statistical feature of course won't give better result compared with the complicated ones, but provide sufficient interpretation of control chart pattern.

This paper discuss about the application of neural networks and correlation statistical feature in practical manner. Both methods need some learning process, i.e. supervising the recognition ability based on training data pattern generated by simulation representing control chart behaviour.

2. Neural networks control chart

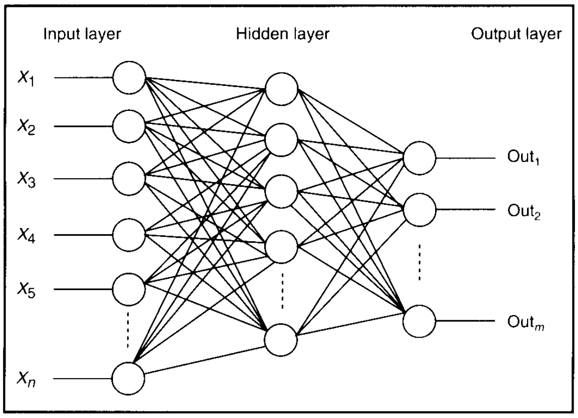

Fig. 1: Neural networks with single hidden layer

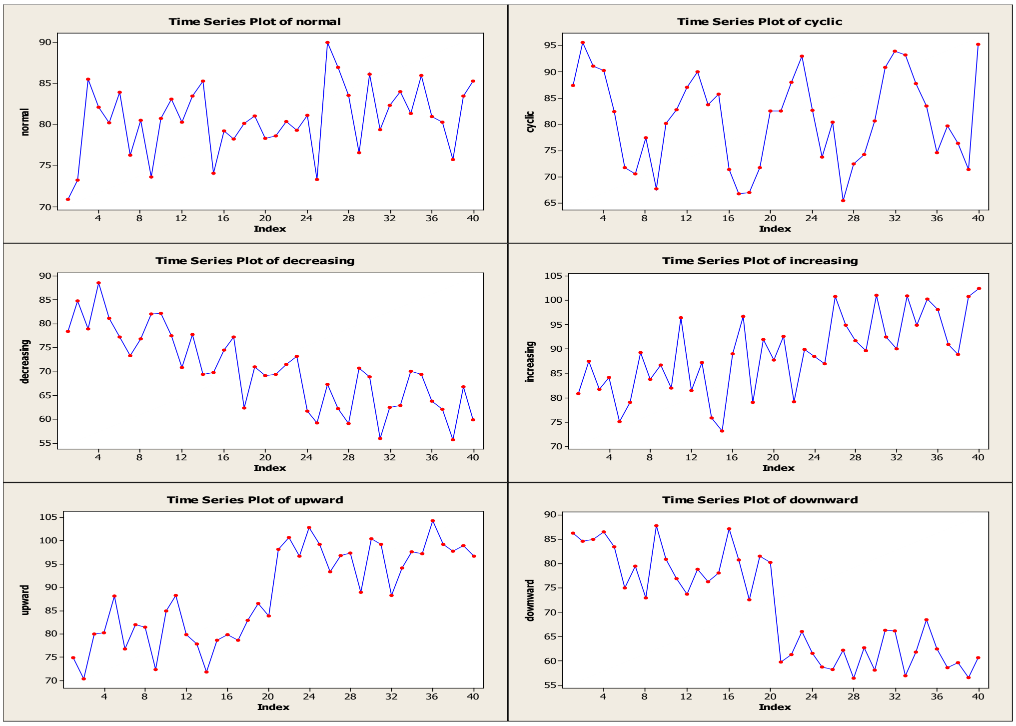

Fig. 2: Control chart patterns

Neural networks consist of input, hidden and output layer. It's a kind of nonlinear mathematical model which can catch any input-output pattern, learn it, and use it as prediction or recognition Bishop [4]. Neurons (represents by circle) in the input layer receive the inputted data and then distributing it to the hidden layer using some activation function. The hidden neurons do the same as the input, receiving information from input layer and then map them to the output layer (if there's only one hidden layer). The output layer then calculates the error component, and then updates the information for previous layers. If the inputted data i.e. the control chart observed data points are represented by $x=(x_1, x_2, \ldots, x_p)$, then the value of each hidden node is:

$$h_j = \phi_j^h\left(w_{oj} + \sum_{i=1}^{p} w_{ij}x_{ij}\right) \tag{1}$$

with parameter (weight) $w_{0j}, \ldots, w_{pj}$ represented by lines connecting between neurons. Logistic activation function $\phi_j^h(.)$ is often used in this architecture. Output value represents the mapping function from the hidden ones:

$$y_i = \phi\left(W_0 + \sum_{j=1}^{r} W_j h_j\right) = \left(W_0 + \sum_{j=1}^{r} W_j \phi\left(w_{0j} + \sum_{i=1}^{p} w_{ij}x_i\right)\right) \tag{2}$$

This recursive mapping algorithm is repeated until a small error bounds has achieved. The single hidden layer neural networks architecture shows in figure 1.

The observed data points on control chart act as data inputted for input layer while the chart pattern translated into coded variables as targeted output layer. Figure 2 shows 6 types of control chart pattern often observed, i.e. normal pattern, cyclic, downward shift, upward shift, increasing trend, and decreasing trend. All these patterns are generated using simulation, and then perform learning process for the neural networks. Equations for simulating control chart patterns as follows [9]:

- Normal patterns:

$$f(t) = r(t)\sigma \tag{3}$$

- Cyclic patterns:

$$f(t) = r(t)\sigma + a \sin(2\pi t/T) \tag{4}$$

- Increasing trend patterns:

$$f(t) = r(t)\sigma + gt \tag{5}$$

- Decreasing trend patterns:

$$f(t) = r(t)\sigma - gt \tag{6}$$

- Upward shift patterns:

$$f(t) = r(t)\sigma + bs \tag{7}$$

- Downward shift patterns:

$$f(t) = r(t)\sigma - bs \tag{8}$$

Function $r(t)$ follows normal distribution with mean μ which can be set to any value, and standard deviation σ. T is the period of cycle in cyclic pattern, g is the gradient of increasing/decreasing pattern, and b indicates the shift position in an upward shift/down pattern. Five kinds of each pattern were simulated as part of neural networks learning process, and as output, the representation code for each pattern in table 1 was used for targeted value as in Sagiroglu et al.[7]

Table 1. Control chart pattern and targeted output for neural networks

Identification	Pattern	Neural networks output					
1	Normal	1	0	0	0	0	0
2	Cyclic	0	1	0	0	0	0
3	Increasing trend	0	0	1	0	0	0
4	Decreasing trend	0	0	0	1	0	0
5	Upward shift	0	0	0	0	1	0
6	Downward shift	0	0	0	0	0	1

Using standard backpropagation algorithm in neural networks, the learning process was performed with thirty kinds of pattern and desired output code. Some neural networks architecture can be chosen for this process, but in this research the single layer neural networks with small number of hidden neuron gives the best result.

3. Correlation based recognition control chart

Yang and Yang [9] proposed control chart pattern recognition based on correlation coefficient. This method needs no sophisticated algorithm in the recognition process, only by calculating the correlation coefficient between observed control chart and the simulation ones, i.e. r_{normal}, r_{cyclic}, $r_{increase}$, $r_{decrease}$, r_{upward}, and $r_{downward}$. The highest correlation coefficient *r*, relating to each simulated patterns means that the observed pattern is closer to one of them. This process represents the learning process as in the neural networks control chart, but gives simpler approach.

4. Result and discussion

Some neural networks architectures were used for modelling, but one with five neurons in hidden layer gives best result. This research hasn't been focusing with searching the best neural networks architectures as done by [5], but only for applying the neural networks and correlation based control chart for real data.

The neural networks learning process was done with 17 backpropagation iterations and targeted error 0.01. Weights estimated in the learning process then are used to predict new cases i.e. observed control chart with some pattern to be recognize. Convergence epoch was reached, and gives 100% accurate recognition for training-data. Generally, steps for performing neural networks control chart are described as follows:

Step 1: generate/simulate the patterns of control chart, using equation (3) to (8)

Step 2: perform neural networks training process with particular architecture, results the estimated weights

Step 3: use the estimated weight due to the architecture to predict or recognize the pattern of new real data

Step 4: read the classification provided by neural networks, i.e. prediction of control chart pattern as coded in Table 1.

As the neural networks control chart, the correlation based control chart needs steps in applying it with little difference in steps order, i.e.:

Step 1: generate/simulate the patterns of control chart, using equation (3) to (8)

Step 2: calculate the correlation coefficient between the simulated ones and the new real data

Step 3: find the highest corresponding correlation coefficient, where the most similar pattern is represented by.

The results were then evaluated due to the best recognition between both methods.

A case study taken from real time observation was used to implements both methods. Data was recorded in a battery factory, observing the internal resistant in battery measured in ohm. Forty points of observation were taken daily, repeated for eight days and charted as in figure 3. Visually, all these observed data was looked like normal patterns, there were no indication for unusual patterns. In order to equalize the measurement unit between simulated and observed data, all of them were normalized before the learning process begun.

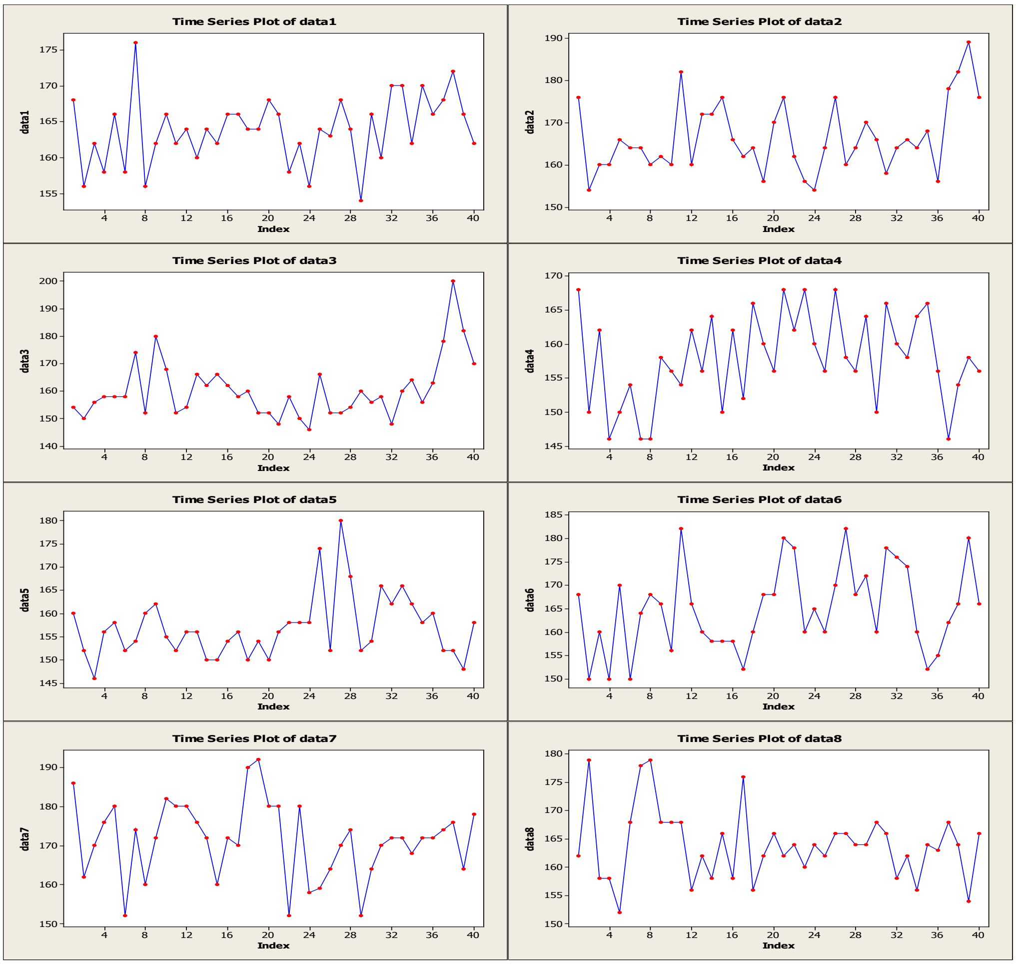

Fig. 3: Observed real battery data

Table 2. Neural networks control chart recognition result

observed data	Neural networks control chart recognition					
	normal	cyclic	incre	decre	upward	downward
day 1	1	0	0	0	0	0
day 2	0	1	-1	1	1	1
day 3	-1	-1	1	1	1	1
day 4	1	0	0	0	0	0
day 5	0	1	0	0	0	0
day 6	0	0	-1	0	1	0
day 7	1	0	1	1	0	1
day 8	-1	0	1	-1	0	1

Table 3. Correlation coefficient from correlation based recognition control chart

observed data	correlation based recognition control chart					
	normal	cyclic	incre	decre	upward	downward
day 1	0.05	0.1	0.1	0.1	0.07	0.01
day 2	0.07	0.02	0.15	-0.1	0.1	-0.1
day 3	0.2	0.08	0.03	0.02	0.1	0.1
day 4	-0.02	0.04	0.1	0.12	0.2	0.11
day 5	0.08	0.05	0.3	-0.3	0.4	-0.3
day 6	0.1	0.15	-0.03	0.06	0.35	-0.2
day 7	0.3	0.1	0.1	-0.1	0.02	0.1
day 8	0.04	0.02	0.01	0.01	-0.01	0

Logically, if the both method is applied to those eight normal pattern observed data, the recognition result should be tend to the same pattern. Evaluating the both method, table 2 and 3 gives their performance. From 8 days, the neural networks control chart only recognizes 2 days (day 1 and day 4) for normal pattern. For other days, no correct classification was performed by this method. Regardless from too few simulating pattern, the neural networks control chart gives unsatisfactory result.

Similar to the neural networks control chart, the correlation based recognition also gives no better recognition result. Only day 3 and day 7 could be correctly classified with the highest correlation coefficient. However, this method gives simpler approach with faster recognition process rather than neural networks one.

5. Conclusion

Both methods in recognizing the control chart pattern, didn't give superior result. But, as the simple one, the correlation method needs no sophisticated algorithm to do the recognition process. This statistical method still can be an alternatives to the artificial intelligences.

For further research, simulated patterns provided for training process must be increased in case of its numbers of various behaviours. The $r(t)$ also can be considered to accommodate various mean and variances, as well the constant b, g, T and s. More number of training data, means that the learning process catches the behaviour of control chart.

6. References

[1] Al-Ghanim, A., (1997), "An Unsupervised Learning Neural Algorithm for Identifying Process Behavior on Control Chart and A Comparison With Supervised Learning", *Computers and Industrial Engineering*, Vol. 32, No. 3, pp 627-639

[2] Anagun, A.S., (1998), "A Neural Network Applied to Pattern Recognition in Statistical Process Control", *Computers and Industrial Engineering*, Vol. 35, No. 1-2, pp 185-188

[3] Ebrahimzadeh, A., and Ranaee, V., (2010), "Control Chart Pattern Recognition using Optimized Neural Network and Efficient Features", *ISA Transactions*, Vol. 49, Issue 3, pp 387-393

[4] Bishop, C.M., (1999), *Neural Networks for Pattern Recognition*, Oxford Univ Press, NY.

[5] Hassan, A., Baksh, S.N., Shaharoun, A.M., Jamaluddin, H., (2003), "Improved SPC Chart Pattern Recognition Using Statistical Features", *International Journal of Production Research*, Vol. 41., No. 7, pp 1587-1603

[6] Masood, I., and Hassan, A., (2010), "Issues in Development of Artificial Neural Network Based Control Chart Pattern Recognition Schemes", *European Journal of Scientific Research*, Vol. 39, No. 3, pp 336-355

[7] Sagiroglu, S., Besdok, E., Erler, M., (2000), "Control Chart Pattern Recognition Using Artificial Neural Networks, *Turkey Journal of Electrical Engineering*, Vol. 8, No. 2, pp 137-147

[8] Wiel, A.A.V., Tucker, W.T., Faltin, F.W., Doganaksoy, N., (1992), "Algorithmic Statistical Process Control: Concept and Applications", *Technometrics*, Vol. 34, no. 3, pp 286-297

[9] Yang, J.H., and Yang, M.H., (2005), "A Control Chart Pattern Recognition System Using A Statistical Correlation Coefficient Method", *Computers and Industrial Engineering*, Vol. 48, pp 205-221