mathematics

MDPI

Article

Using Machine Learning Classifiers to Recognize the Mixture Control Chart Patterns for a Multiple-Input Multiple-Output Process

Yuehjen E. Shao * and Yu-Ting Hu

Department of Statistics and Information Science, Fu Jen Catholic University, Xinzhuang Dist., New Taipei City 24205, Taiwan; 404336075@mail.fju.edu.tw
* Correspondence: stat1003@mail.fju.edu.tw

Received: 29 November 2019; Accepted: 6 January 2020; Published: 7 January 2020

check for updates

Abstract: A statistical process control (SPC) chart is one of the most important techniques for monitoring a process. Typically, a certain root cause or a disturbance in a process would result in the presence of a systematic control chart pattern (CCP). Consequently, the effective recognition of CCPs has received considerable attention in recent years for their potential use in improving process quality. However, most studies have focused on the recognition of CCPs for SPC applications alone. Specifically, even though numerous studies have addressed the increased use of the SPC and engineering process control (EPC) mechanisms, very little research has discussed the recognition of CCPs for multiple-input multiple-output (MIMO) systems. It is much more difficult to recognize the CCPs of an MIMO system since two or more disturbances are simultaneously involved in the process. The purpose of this study is thus to propose several machine learning (ML) classifiers to overcome the difficulties in recognizing CCPs in MIMO systems. Because of their efficient and fast algorithms and effective classification performance, the considered ML classifiers include an artificial neural network (ANN), support vector machine (SVM), extreme learning machine (ELM), and multivariate adaptive regression splines (MARS). Furthermore, one problem may arise due to the existence of embedded mixture CCPs (MCCPs) in MIMO systems. In contrast to using typical process outputs alone in a classifier, this study employs both process outputs and EPC compensation to ensure the effectiveness of CCP recognition. Experimental results reveal that the proposed classifiers are able to effectively recognize MCCPs for MIMO systems.

Keywords: control chart pattern; SPC; EPC; MIMO; artificial neural network; support vector machine; extreme learning machine; multivariate adaptive regression splines

1. Introduction

Because of their ability to detect disturbances, statistical process control (SPC) charts are widely used in monitoring industrial processes. A process is considered to be out of control when systematic patterns are exhibited in SPC charts [1,2]. Disturbances contribute to the presence of systematic control chart patterns (CCPs) for a process. The recognition of CCPs is very important since CCPs are typically associated with certain root causes that antagonistically influence the process [3–6].

Because it is important for process personnel to determine root causes for process improvement, many studies have discussed the effectiveness of CCP recognition through various machine learning (ML) mechanisms. An artificial neural network (ANN) approach was discussed to recognize CCPs for a multivariate process [7]. ANNs were employed to identify a set of subclasses of abnormal multivariate CCPs, and the χ^2 statistic served as the input to the ANNs. Additionally, the proposed mechanism was evaluated for a real case study, and good results were reported. In [8], a hybrid

ANN model was proposed to detect and classify single and mixture CCPs (MCCPs), determine the major corresponding parameters and estimate the starting point of unnatural patterns. That work also developed a method for pattern simulation in variable and attribute control charts.

In [9], researchers indicated that ANNs with features extracted from process data as an input vector representation can facilitate efficient pattern recognition. Based on a set of seven shape features, they designed a multilayered perceptron neural network to recognize CCPs for a process. A procedure, including a clustering module and a classifier module, was proposed to recognize CCPs [10]. The classifier module, which contained the multilayer perceptron, probabilistic neural networks, and radial basis function neural networks (RBFNNs), was used to determine the membership of the patterns. In [11], a mechanism based on an RBFNN was proposed for CCP recognition. Four modules, including feature extraction, feature selection, classification, and a learning algorithm, were designed in the proposed mechanism. The experimental results demonstrated that the proposed approach had good performance for recognition of eight CCPs. The researchers also designed two modules, a feature extraction module and a classifier module, for CCP recognition. Multilayer perceptron neural networks and a radial basis function were used in the classifier module [12].

In addition to the use of ANNs, support vector machine (SVM) is another commonly used classifier for CCP recognition. A method combining fuzzy SVM, a hybrid kernel function, and a genetic algorithm (GA) was designed for CCP recognition [13]. The results showed that the fuzzy SVM-based classifier outperformed the methods of a learning vector quantization network, multilayer perceptron network, and probability neural network. In [6], a weighted support vector machine (WSVM) method was proposed for automated process monitoring. The researchers discussed the superiority of WSVM over traditional SVM in the situations of various fault scenarios. A hybrid procedure containing independent component analysis (ICA) and SVM was proposed for CCP recognition [3]. In this procedure, ICA was used to capture independent components (ICs), and the ICs served as the input variables of the SVM for building a CCP recognition model. In [14], the researchers indicated that SVM classifiers were better than ANNs for CCP recognition. They also demonstrated an SVM-based CCP recognition model for process outputs with first-order autoregressive (AR(1)) noise. Instead of identifying single CCPs, they designed a hybrid model that included the wavelet transform and improved particle swarm optimization-based support vector machine (P-SVM) for the recognition of MCCPs [15]. A raw mixture pattern signal was decomposed into two basic pattern signals through the wavelet transform. Then, those two basic pattern signals were recognized by multiclass SVMs. Because of the generalization capability, they presented a fuzzy SVM classifier for CCP recognition [16]. The proposed combined scheme included the fuzzy classifier subnetwork, SVM subnetwork, and optimization subnetwork.

While most studies have focused on the use of ANN and SVM classifiers for CCP recognition, extreme learning machine (ELM) techniques have been used to identify MCCPs for a process [17,18]. The multivariate adaptive regression splines (MARS) scheme was also applied to the recognition of MCCPs for an SPC and engineering process control (EPC) process [19,20].

By conducting this little review, we have found that CCP recognition has been extensively addressed; however, little research has investigated CCP recognition for an SPC-EPC system. The reasons why SPC-EPC is important to industrial processes can be described as follows. Traditional SPC charts presume that the process outputs should be independent, which is not a practical consideration for real-world applications [21–23]. Correlated process outputs cause increases in false alarm signals and misinterpretation of the function of SPC charts [4,18]. In fact, correlation broadly exists in continuous and chemical processes [4,24,25]. The integration of EPC and SPC mechanisms is commonly used to overcome this independence difficulty for SPC applications [18,20,26,27]. Although the approach is able to overcome the correlation problems for SPC applications, the use of EPC may conceal the effects of underlying process disturbances. These embedded disturbance effects imply that process personnel have more difficulty recognizing underlying CCPs [28].

In [4], the systems were represented as an AR(1), a moving average (MA(1)) or an autoregressive moving average (ARMA(1, 1)) model. The effect of autocorrelation on the pattern recognition capabilities was investigated. However, the study did not utilize EPC actions for CCP recognition. The authors also reported that the first-order integrated moving average (IMA(1, 1)) component was used as the noise in the SPC-EPC system [19,20]. Although they considered EPC actions in the CCP recognition scheme, they only considered the univariate system of an SPC-EPC. In this study, we refer to a multivariate SPC-EPC system as a multiple-input multiple-output (MIMO) process. It is a challenging task to recognize the underlying CCPs in an MIMO system since two or more process outputs, EPC actions, and various disturbances are involved in the system. Little attention has been given to the utilization of ML classifiers for CCP recognition in MIMO systems that are commonly encountered in practical manufacturing processes. Accordingly, the purpose of the present study is to present an effective approach for recognizing MCCPs for an MIMO system.

Because the ANN and SVM techniques have the ability to seize nonlinear and complex features of an SPC-EPC process with a high degree of accuracy, this study is motivated to propose ANN and SVM classifiers for MCCP recognition for an MIMO system. The present study also considers ELM and MARS as classifiers for the proposed mechanism. The reason for using ELM is that it has the advantages of fast learning speed and good generalization performance [29,30]. The reason for choosing MARS is that it has been adopted for CCP recognition in only a few studies, although MARS is effective in classification [31,32].

Since they are used to tune the MIMO system, EPC actions are extracted and serve as the inputs for the classifiers in this study. Experimental results show that the proposed ML classifiers with EPC inputs are able to effectively recognize various MCCPs for an MIMO system. The rest of this study is organized as follows. Section 2 discusses models of MIMO systems and four types of disturbances. Additionally, the difficulty of MCCP recognition for an MIMO system is presented. Four ML techniques for CCP recognition used in this study are introduced in Section 3. Section 4 presents the results of simulated experiments and demonstrates the performance of the proposed approaches. The final section addresses the research findings and conclusions inferred from this study.

2. MIMO and Disturbance Models

2.1. MIMO Models

Suppose a process with m inputs and p outputs is expressed by [33–35]:

$$y_i \;=\; \alpha + \beta x_{i-1} + d_i, \tag{1}$$

where $y_i(p \times 1)$ represents a vector of process outputs at time i, $\alpha(p \times 1)$ is the offset parameters, $\beta(p \times m)$ is defined as the gain parameters, $x_i(m \times 1)$ is a vector of controllable variables at time i (i.e., EPC adjustments), and $d_i(p \times 1)$ is defined as the process noise at time i. This study assumes that the process noise follows a white noise.

This study assumes that B is the estimate of β. In addition, this study assumes that $\hat{\alpha}_0$ is the estimate of α at $i = 0$. Because the output offset would be updated after each run, the predicted model is depicted as follows:

$$\hat{y}_i = \hat{\alpha}_0 + Bx_{i-1}. \tag{2}$$

Before implementing the EPC adjustments, the initial process input is:

$$x_0 = B^{-1}(\tau - \hat{\alpha}_0), \tag{3}$$

where τ is the target vector.

This study considers a multivariate exponentially weighted moving average (MEWMA) controller as the EPC scheme [33]. This MEWMA controller is defined as follows:

$$\hat{\alpha}_i = \omega\,(y_i - Bx_{i-1}) + (1-\omega)\,\hat{\alpha}_{i-1}, \tag{4}$$

where ω is a discount factor, and the range of ω is from 0 to 1. For the ith run, x_i can be described as follows:

$$x_i = \left(I - B'\left(BB'\right)^{-1}B\right)x_{i-1} + B'\left(BB'\right)^{-1}\left(\tau - \hat{d}_i\right). \tag{5}$$

Let $\alpha_0 = 0$ and $\tau = 0$; then, the off-target amount at run i can be described as follows:

$$y_i = (1-\omega)^{i-1}\gamma_0 + \sum_{t=0}^{i-1}(1-\omega)^t(d_{i-t} - d_{i-t-1}). \tag{6}$$

Since d_i has mean vector μ and variance Σ, the covariance of y_i becomes:

$$\Sigma_{y_i} = (1 + \frac{\omega}{2-\omega}(1-(1-\omega)^{2(i-1)}))\Sigma. \tag{7}$$

2.2. Disturbance Models

For an MIMO process, disturbances may interrupt the system at any time. When a disturbance has intruded, the noise for an MIMO process can be expressed as follows:

$$d_i = D_i, \tag{8}$$

where D_i stands for a vector of certain disturbance at time i. As a result, Equation (1) is remodeled as follows:

$$y_i = \alpha + \beta x_{i-1} + D_i. \tag{9}$$

This study examines four types of disturbances for an MIMO process. Those disturbances are defined as follows [15,36]:

$$\text{Cycle}: D_i^{CYC} = \sin\left(\frac{2\pi i}{e}\right)A_i + a_i, \tag{10}$$

$$\text{Systematic}: D_i^{SYS} = h(-1)^i + a_i, \tag{11}$$

$$\text{Shift}: D_i^{SHI} = R_i + a_i, \tag{12}$$

$$\text{Stratification}: D_i^{STA} = va_i, \tag{13}$$

where D_i^{CYC} represents the value of cycle (CYC) disturbance at time i, A_i stands for the cycle amplitude, which is assumed to follow a uniform distribution within the range of (1.5, 2.5), e represents the cycle period, which is assumed to be e = 8, D_i^{SYS} represents the value of systematic (SYS) disturbance at time i, and h stands for the magnitude of the systematic pattern in terms of σ^2, which is assumed to follow a uniform distribution within the range of (1.0, 3.0), D_i^{SHI} represents the value of shift (SHI) disturbance at time i, R_i stands for the level of shift disturbance, which is assumed to follow a uniform distribution within the range of (1, 2), D_i^{STA} represents the value of stratification (STA) disturbance at time i, and v stands for random noise, which is assumed to follow a uniform distribution within the range of (0.2, 0.4).

3. Methodology and Experimental Results

In [19], although their CCP recognition designs were associated with EPC actions, they only implemented their designs for a univariate SPC-EPC system. However, it is much more complex to recognize MCCPs for MIMO systems since more than two process outputs and EPC actions are involved. Furthermore, in addition to the method of ELM, this study will employ their methods (i.e., ANN, SVM and MARS) to illustrate the classification capability.

Because the ANN, SVM, and MARS techniques were introduced by [19,37–39], this study briefly addresses the modeling concepts of ELM here.

3.1. ELM Classifier

For ELM modeling, it randomly assigns the input weights and analytically determines the output weights for the single-hidden-layer feedforward neural networks (SFNNs) [40]. The concept of ELM is provided as follows [40]. Suppose there are N arbitrary distinct samples (x_i, t_i), where $x_i = [x_{i1}, x_{i2}, \ldots, x_{in}]^T \in R^n$ and $t_i = [t_{i1}, t_{i2}, \ldots, t_{im}]^T \in R^m$. SFNNs with $\widetilde{N}$ hidden neurons and activation function $f(x)$ can reach N samples with zero error. Thus [40],

$$F\beta = T, \tag{14}$$

where

$$F\left(w_1, \ldots, w_{\widetilde{N}}, b_1, \ldots, b_{\widetilde{N}}, x_1, \ldots, x_N\right) = \begin{bmatrix} f(w_1 \cdot x_1 + b_1) & \cdots & f\left(w_{\widetilde{N}} \cdot x_1 + b_{\widetilde{N}}\right) \\ \vdots & \ddots & \vdots \\ f(w_1 \cdot x_N + b_1) & \cdots & f\left(w_{\widetilde{N}} \cdot x_N + b_{\widetilde{N}}\right) \end{bmatrix}_{N \times \widetilde{N}};$$

$$\beta_{\widetilde{N} \times m} = (\beta_1^T, \ldots, \beta_{\widetilde{N}}^T)^t;$$

$$T_{N \times m} = (T_1^T, \ldots, T_N^T)^t;$$

$$w_i = [w_{i1}, w_{i2}, \ldots, w_{in}]^T;$$

$i = 1, 2, \ldots, \widetilde{N}$, represents the weight vector connecting the ith hidden node and the input nodes; $\beta_i = [\beta_{i1}, \beta_{i2}, \ldots, \beta_{im}]^T$, stands for the weight vector connecting the ith hidden node and the output nodes; $w_i \cdot x_j$ stands for the inner product of w_i and x_j; b_i stands for the threshold of the ith hidden node; and F represents the hidden layer output matrix.

Subsequently, the establishment of the output weights attain the least-square solution of the given linear system. The minimum norm least-square solution of the linear system becomes

$$\hat{\beta} = F^{\Psi} T, \tag{15}$$

where F^{Ψ} is the Moore–Penrose generalized inverse of matrix F.

3.2. MCCPs for an MIMO Process

This study assumes that an MIMO process is disturbed by four single disturbances, which are described by Equations (10)–(13). When any one of those four disturbances (i.e., Equations (10)–(13)) is introduced in the MIMO system, this study refers to that situation as the presence of a single CCP. When any two or more of those four disturbances are concurrently present in the MIMO system, we refer to that situation as the presence of an MCCP in this study. Additionally, this study considers two cases of process outputs for an MIMO process. The first case is that two process outputs and two single disturbances concurrently exist in an MIMO process. The second case is that three process outputs and three single disturbances concurrently exist in an MIMO process. Since this study assumes that two and three single disturbances may be concurrently introduced into the MIMO process, there are six and four types of MCCPs that need to be recognized. Table 1 shows those types of MCCPs for two and three process outputs, respectively.

To demonstrate the difficulty of recognizing MCCPs for an MIMO process, this study performs a series of computer simulations. Assume that an MIMO process (i.e., Equation (1)) includes two process outputs (i.e., y_1 and y_2). Additionally, suppose that there are no disturbances occurring before time 100, and the mean vector of the process output is on target at $[0, 0]^1$ for the first 100 observations. After time

101, single SHI and CYC disturbances are concurrently introduced into the first and second outputs of the MIMO process, respectively. Figure 1 shows the output observations of the MIMO system for 600 runs. Figure 1 clearly shows that the SHI and CYC disturbances are hardly differentiated since the two outputs exhibit similar patterns. A similar pattern results from the compensation of EPC. Figure 2 displays the three output observations when single SHI, CYC, and STR disturbances are concurrently introduced into an MIMO process after time 101. These three disturbances are much more difficult to differentiate.

Table 1. Types of mixture control chart patterns (MCCPs) for two and three single disturbances in an MIMO process.

2 Single Disturbances	3 Single Disturbances
(1) {SHI-CYC}	(1) {SHI-CYC-SYS}
(2) {SHI-SYS}	(2) {SHI-CYC-STR}
(3) {SHI-STR}	(3) {SHI-SYS-STR}
(4) {CYC-SYS}	(4) {CYC-SYS-STR}
(5) {CYC-STR}	
(6) {SYS-STR}	

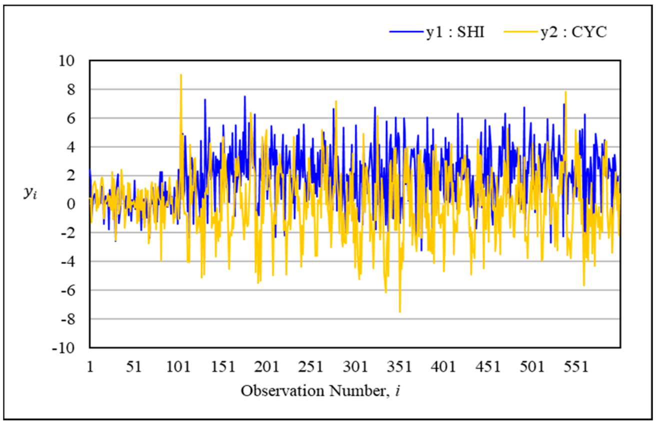

Figure 1. Two multiple-input multiple-output (MIMO) process outputs, which concurrently contain single SHI and CYC disturbances after time 100.

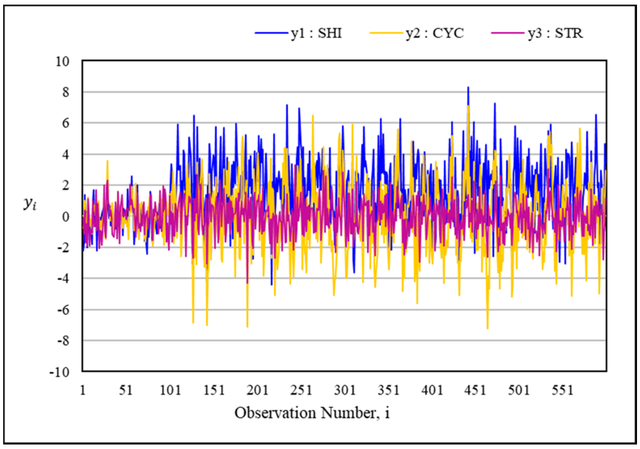

Figure 2. Three MIMO process outputs, which concurrently contain single SHI, CYC and STR disturbances after time 100.

3.3. Research Framework

Figure 3 presents a generalized depiction of the research methodologies used in this study. As shown in Figure 3, from left to right, the data vectors for the classifiers are generated by computer simulations. In this study, RStudio [41] was used to generate the MIMO process data vectors. Additionally, because the previous studies used the ratio of 7:3 for training and testing data vectors [19,20], we followed their suggestion. Therefore, a ratio of 7:3 for training and testing data vectors is used for all cases. The four ML classifiers, ANN, SVM, ELM, and MARS, are investigated for MCCP recognition. The final phase addresses the recognition results and comparative analysis. To recognize the MCCPs for an MIMO system, we design two different schemes for those four classifiers. In the first scheme, the process outputs y (i.e., Equation (6)) is the only variable that serves as the input for the classifiers. Therefore, the first scheme is similar to traditional approaches for CCP recognition. Since this study considers the case of EPC in the underlying system, the second scheme includes x (i.e., Equation (5)) and y (i.e., Equation (6)) as the classifiers' inputs and considers Z (i.e., the classification category) as the classifiers' output.

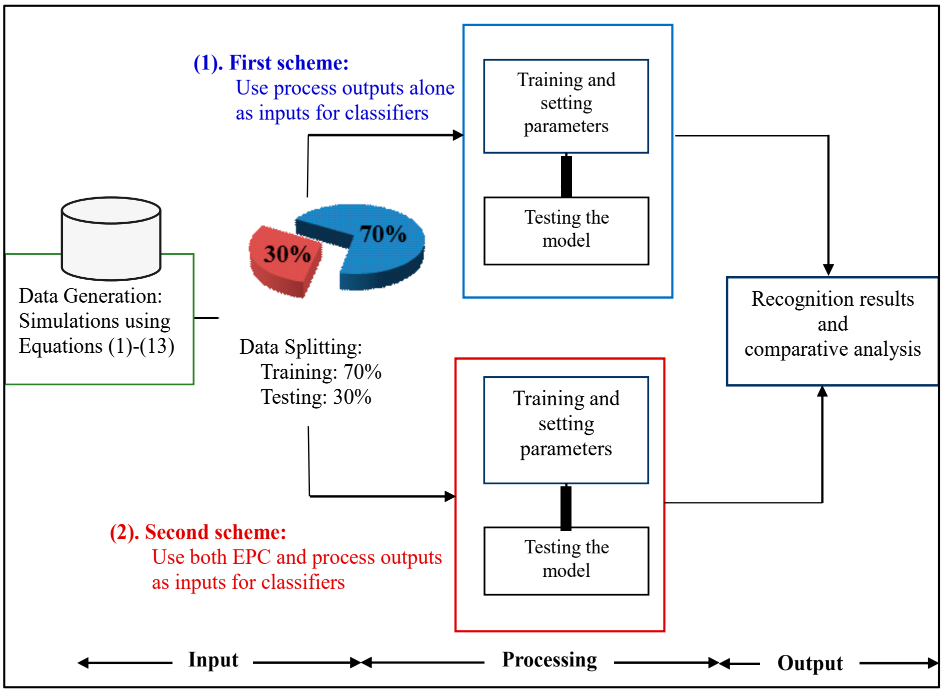
(1). First scheme:
Use process outputs alone as inputs for classifiers
Training and setting parameters
Testing the model
30%
70%
Data Generation: Simulations using Equations (1)-(13)
Data Splitting: Training: 70% Testing: 30%
Recognition results and comparative analysis
Training and setting parameters
Testing the model
(2). Second scheme:
Use both EPC and process outputs as inputs for classifiers
Input
Processing
Output

Figure 3. The research framework.

3.4. Four Cases and the Recognition Results

(1) Case 1: Two process outputs with the first scheme (i.e., without the use of EPC)

In this study, the first scheme only employs the process outputs y to serve as the input variable for the classifiers. In the case of two outputs for an MIMO process, this study uses 1400 and 600 data vectors for the training and testing phases, respectively. As an example, by considering the type of {SHI-CYC} in Table 1, the first 700 data vectors are generated from the first process output (i.e., y_1) with the presence of {SHI} alone, and the data vectors from 701 to 1400 are generated from the second process output (i.e., y_2) with the presence of {CYC} alone. The testing data structure is similar to the training data structure. That is, the first 300 data vectors are generated from the first process output (i.e., y_1) using the {SHI} disturbance alone, and the final data vectors from 301 to 600 are generated from the second process output (i.e., y_2) using the {CYC} disturbance alone. In this example, this study sets the values of the output node (i.e., Z) as either 0 or 1. A value of 0 represents the presence of the SHI disturbance, and a value of 1 stands for the presence of the CYC disturbance.

This study considers the accurate identification rate (AIR) as the model performance measurement. The AIR is described as follows:

$$\text{AIR} = \frac{n_a}{N}, \tag{16}$$

where N represents the total number of data vectors used for the recognition process and n_a stands for the number of data vectors in N where the true disturbance type is accurately recognized.

After performing classification tasks with the ANN, SVM, ELM and MARS classifiers, Table 2 shows the recognition results for six types of MCCPs for an MIMO process. In this study, the parameters of ANN include $\{n_i, n_h, n_o\}$. The meanings of those parameters are the number of neurons in the input layer, hidden layer and output layer, respectively. Typically, because there are no specific rules for choosing the number of hidden nodes, we use the rules of thumb to determine the numbers for the ANN design. The hidden nodes were settled to range from $(2a - 2)$ to $(2a + 2)$, where a is the number of input variables. Because $a = 1$ in this experiment, the hidden nodes were settled to 1, 2, 3, or 4. Furthermore, we use the learning rate at the default value (i.e., 0.01) to ensure consistency. A total of 10,000 training iterations (i.e., default value) were implemented for each network. The ANN design with the highest AIR is employed as the optimal network topology.

Table 2. Results of MCCP recognition for two process outputs (first scheme: without the use of EPC).

Types of MCCPs	AIR-ANN (Parameters: $\{n_i, n_h, n_o\}$)	AIR-SVM (Parameters: $\{C, \gamma\}$)	AIR-ELM (Parameters: $\{E_n, E_h, E_o\}$)	AIR-MARS (Parameters: {null})
SHI-CYC	68.50% {1, 1, 1}	68.67% $\{2^{-5}, 2^{-5}\}$	68.60% {1, 6, 1}	68.50%
SHI-SYS	65.00% {1, 4, 1}	64.00% $\{2^{-5}, 2^{-5}\}$	65.32% {1, 6, 1}	64.17%
SHI-STR	76.50% {1, 3, 1}	77.50% $\{2^{-4}, 2^{-5}\}$	77.05% {1, 5, 1}	76.50%
CYC-SYS	51.50% {1, 4, 1}	57.33% $\{2^{-2}, 2^{-5}\}$	56.38% {1, 29, 1}	58.50%
CYC-STR	64.33% {1, 3, 1}	64.17% $\{2^{-1}, 2^{-5}\}$	61.00% {1, 6, 1}	64.83%
SYS-STR	73.67% {1, 2, 1}	73.00% $\{2^{0}, 2^{-5}\}$	71.81% {1, 18, 1}	73.00%

For the SVM design, because the radial basis function (RBF) is one of the most popular kernel functions for various learning algorithms, we use an RBF in our experiments. In addition, the values of two parameters (C and γ) would mainly affect the performance of the SVM [38]. There are no specific rules for determining the values of C and γ. This study uses a grid search for the parameter settings [42]. The grid search method utilizes exponentially growing sequences of C and γ to obtain the appropriate parameters. The settings of C and γ that would generally generate the highest AIR were selected. The trained SVM model with the best parameter settings, denoted $\{C, \gamma\}$, is employed for the MCCP recognition process.

For the ELM design, we used the notation of $\{E_n, E_h, E_o\}$ to represent the parameter settings. They include the number of neurons in the input layer, hidden layer and output layer, respectively. The most important ELM parameter is the number of hidden nodes. Because the previous study used the hidden nodes varying from 1 to 30 [19], we followed their suggestion. For each number of nodes, an ELM model is repeated 30 times, and the average root mean squared error (RMSE) of each node is computed. The number of hidden nodes that provides the smallest average RMSE value is chosen as the best parameter setting for the ELM scheme. Furthermore, because there are no specific parameter settings for the MARS scheme, this study simply denotes the parameter settings as {null} for the MARS classifier.

Table 2 shows that the average AIRs of MCCPs are 66.58%, 67.45%, 66.69%, and 67.58% for the ANN, SVM, ELM, and MARS classifiers, respectively. The accurate identification percentages for MCCP recognition seem to be acceptable but are not satisfactory. Accordingly, the second scheme of the classifiers is considered for performing MCCP recognition tasks.

(2) Case 2: Two process outputs with the second scheme (i.e., with the use of EPC)

In this study, the second scheme employs both EPC actions x and process outputs y as the input variables for the classifiers. For this case, this study uses two sets of inputs for the training and testing phases of the classifiers.

Again, consider the type of {SHI-CYC} in Table 1 as an example. The first set of inputs is the EPC actions, and the first 700 data vectors are captured from the first EPC outputs (i.e., x_1) with the compensation of {SHI}, and the data vectors from 701 to 1400 are captured from the second EPC outputs (i.e., x_2) with the compensation of {CYC}. The second set of inputs is the process outputs; the first 700 data vectors are generated from the first process outputs (i.e., y_1) with the presence of {SHI}, and the data vectors from 701 to 1400 are generated from the second process outputs (i.e., y_2) with the presence of {CYC}.

The testing data structure is as follows. The first input set contains 600 data vectors. The first 300 data vectors contain the first EPC outputs (i.e., x_1) with the compensation of {SHI}, and the data vectors from 301 to 600 are generated from the second EPC outputs (i.e., x_2) with the compensation of {CYC}. The second input set also contains 600 data vectors. The first 300 data vectors contain the first process outputs (i.e., y_1) with the compensation of {SHI}, and the data vectors from 301 to 600 are generated from the second process outputs (i.e., y_2) with the compensation of {CYC}. The structure of the output node of the classifiers is the same as in case 1. That is, a value of $Z = 0$ represents the presence of the SHI disturbance, and a value of $Z = 1$ stands for the presence of the CYC disturbance.

After implementing classification tasks with the ANN, SVM, ELM, and MARS classifiers using both EPC and process outputs as inputs, Table 3 displays the recognition results for six types of MCCPs for an MIMO process. Table 3 shows that the average AIRs of MCCPs are 81.09%, 82.72%, 81.41%, and 78.19% for the ANN, SVM, ELM, and MARS classifiers, respectively. The accurate identification percentages for MCCP recognition are now satisfactory.

Table 3. Results of MCCP recognition for two process outputs (second scheme: with the use of EPC).

Types of MCCPs	**AIR$_{\text{-ANN}}$** **(Parameters: {n_i, n_h, n_o})**	**AIR$_{\text{-SVM}}$** **(Parameters: {C, γ})**	**AIR$_{\text{-ELM}}$** **(Parameters: {E_n, E_h, E_o})**	**AIR$_{\text{-MARS}}$** **(Parameters: {null})**
SHI-CYC	99.67% {2, 3, 1}	99.67% {2^5, 2^0}	99.93% {2, 29, 1}	99.17%
SHI-SYS	100.00% {2, 2, 1}	100.00% {2^3, 2^5}	99.98% {2, 30, 1}	100.00%
SHI-STR	98.67% {2, 3, 1}	99.83% {2^5, 2^2}	99.76% {2, 30, 1}	99.83%
CYC-SYS	58.67% {2, 6, 1}	53.33% {2^3, 2^{-4}}	58.85% {2, 21, 1}	55.50%
CYC-STR	72.33% {2, 3, 1}	69.83% {2^{-3}, 2^{-5}}	70.57% {2, 29, 1}	62.33%
SYS-STR	57.17% {2, 2, 1}	73.67% {2^{-3}, 2^{-3}}	59.38% {2, 30, 1}	52.33%

(3) Case 3: Three process outputs with the first scheme (i.e., without the use of EPC)

In this case, we use 2100 and 900 data vectors for the training and testing phases, respectively. Since there are three process outputs for an MIMO process, the training phase contains the first 700 data vectors generated from the first process output (i.e., y_1), the data vectors from 701 to 1400 are generated from the second process output (i.e., y_2), and the data vectors from 1401 to 2100 are generated from the third process output (i.e., y_3). The testing phase contains the first 300 data vectors generated from the first process output (i.e., y_1), the data vectors from 301 to 600 are generated from the second process output (i.e., y_2), and the data vectors from 601 to 900 are generated from the third process output (i.e., y_3). In addition, this study sets the values of the output node (i.e., Z) as 0, 1, or 2. A value of 0 represents the presence of the first single disturbance, a value of 1 represents the presence of the second single disturbance, and a value of 2 represents the presence of the third single disturbance. All other parameter settings are the same as in case (1).

After completing classification tasks with the four classifiers, Table 4 shows the recognition results for four types of MCCPs for an MIMO process. This table shows that the average AIRs of MCCPs are 50.42%, 45.00%, 49.44%, and 36.75% for the ANN, SVM, ELM, and MARS classifiers, respectively. Most of the average AIRs are lower than 50.00%, especially for MARS, which has the lowest average AIR value. The accurate identification percentages for MCCP recognition are still not satisfactory. Therefore, this study investigates the use of the second scheme for the four classifiers.

Table 4. Results of MCCP recognition for three process outputs (first scheme: without the use of EPC).

Types of MCCPs	AIR$_{\text{-ANN}}$ (Parameters: $\{n_i, n_h, n_o\}$)	AIR$_{\text{-SVM}}$ (Parameters: $\{C, \gamma\}$)	AIR$_{\text{-ELM}}$ (Parameters: $\{E_n, E_h, E_o\}$)	AIR$_{\text{-MARS}}$ (Parameters: {null})
{SHI-CYC-SYS}	47.11% {1, 1, 1}	42.33% $\{2^5, 2^{-5}\}$	32.23% {1, 10, 1}	33.33%
{SHI-CYC-STR}	53.89% {1, 4, 1}	30.67% $\{2^5, 2^{-5}\}$	48.18% {1, 13, 1}	42.33%
{SHI-SYS-STR}	54.78% {1, 4, 1}	43.00% $\{2^5, 2^{-5}\}$	56.88% {1, 7, 1}	38.22%
{CYC-SYS-STR}	45.89% {1, 3, 1}	64.00% $\{2^5, 2^{-5}\}$	60.45% {1, 14, 1}	33.11%

(4) Case 4: Three process outputs with the second scheme (i.e., with the use of EPC)

In this case, the structure of the data vectors for the classification tasks is almost the same as in case (3), except that one more data vector, the EPC outputs, was included and acted as another input variable for the classifiers. Table 5 shows the outputs of MCCP recognition under case 4. The average AIRs of MCCPs are 64.06%, 67.69%, 52.37%, and 55.84% for the ANN, SVM, ELM, and MARS classifiers, respectively. These AIRs are greatly enhanced after using the extra EPC outputs.

Table 5. Results of MCCP recognition for three process outputs (second scheme: with the use of EPC).

Types of MCCPs	AIR$_{\text{-ANN}}$ (Parameters: $\{n_i, n_h, n_o\}$)	AIR$_{\text{-SVM}}$ (Parameters: $\{C, \gamma\}$)	AIR$_{\text{-ELM}}$ (Parameters: $\{E_n, E_h, E_o\}$)	AIR$_{\text{-MARS}}$ (Parameters: {null})
{SHI-CYC-SYS}	72.11% {2, 3, 1}	67.44% $\{2^5, 2^2\}$	50.00% {2, 29, 1}	66.67%
{SHI-CYC-STR}	77.78% {2, 3, 1}	79.56% $\{2^{-1}, 2^5\}$	60.66% {2, 30, 1}	66.67%
{SHI-SYS-STR}	78.56% {2, 3, 1}	79.33% $\{2^5, 2^3\}$	49.97% {2, 30, 1}	66.33%
{CYC-SYS-STR}	27.78% {2, 6, 1}	44.44% $\{2^5, 2^{-3}\}$	48.86% {2, 30, 1}	23.67%

4. Discussion

In this study, an MIMO system is assumed to be disturbed by four types of disturbances characterized by Equations (10)–(13). This study used ANN, SM, ELM and MARS to recognize the mixture CCPs in an MIMO system.

Because multivariate EPC adjustments were implemented, the disturbances could be embedded in the MIMO system. As a result, the recognition capability of CCPs may not be acceptable. On the other hand, the lack of acceptable recognition capability leads to involvement of some other important factors. Since the recognition performance of the first design is not satisfactory for the cases of two and three process outputs in an MIMO process, this study proposes the use of EPC outputs to serve as an extra input variable for the classifiers. This favorable design brings good improvement in MCCP recognition in an MIMO process.

For MCCP recognition by the ANN classifier, the AIR percentage improvement (denoted by AIR_ANN_{PI}) of the proposed second scheme over the first scheme is defined as follows:

$$AIR\_ANN_{PI} = \frac{(AIR\_ANN_{2nd} - AIR\_ANN_{1st})}{AIR\_ANN_{1st}} * 100\%, \tag{17}$$

where AIR_ANN_{2nd} is the AIR from performing the ANN classifier with the second scheme and AIR_ANN_{1st} is the AIR from performing the ANN classifier with the first scheme.

Similarly, for MCCP recognition by the SVM, ELM and MARS classifiers, the AIR percentage improvements (denoted by AIR_SVM_{PI}, AIR_ELM_{PI}, and AIR_MARS_{PI}, respectively) of the proposed second scheme over the first scheme can be defined as follows:

$$AIR\_SVM_{PI} = \frac{(AIR\_SVM_{2nd} - AIR\_SVM_{1st})}{AIR\_SVM_{1st}} * 100\%, \tag{18}$$

$$AIR\_ELM_{PI} = \frac{(AIR\_ELM_{2nd} - AIR\_ELM_{1st})}{AIR\_ELM_{1st}} * 100\%, \tag{19}$$

$$AIR\_MARS_{PI} = \frac{(AIR\_MARS_{2nd} - AIR\_MARS_{1st})}{AIR\_MARS_{1st}} * 100\%. \tag{20}$$

For MCCP recognition in an MIMO process with two outputs using the ANN classifier, AIR_ANN_{PI} of the second scheme over the first scheme was 21.78%. In addition, AIR_SVM_{PI}, AIR_ELM_{PI}, and AIR_MARS_{PI} of the second scheme over the first scheme were 22.65%, 22.07%, and 15.70%, respectively. Furthermore, for an MIMO process with three outputs using the ANN, SVM, ELM, and MARS classifiers, the associated AIR_ANN_{PI}, AIR_SVM_{PI}, AIR_ELM_{PI} and AIR_MARS_{PI} were 27.05%, 50.43% and 5.94% and 51.94%, respectively.

Figures 4 and 5 display the AIR percentage improvements obtained by employing the proposed scheme instead of the first scheme for an MIMO process with two and three outputs, respectively. As shown in Figures 4 and 5, considerable accuracy improvements can be achieved by using the proposed scheme.

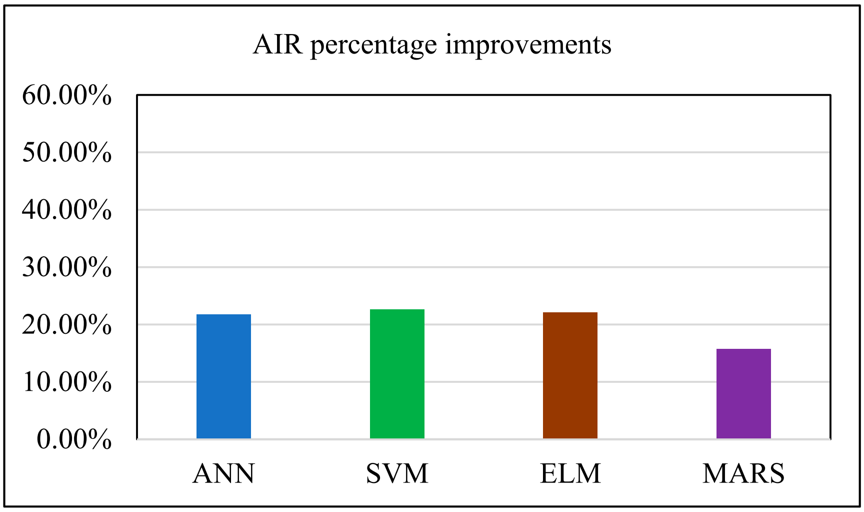

Figure 4. Accurate identification rate (AIR) percentage improvements for an multiple-input multiple-output (MIMO) process with two outputs, using the second scheme instead of the first scheme.

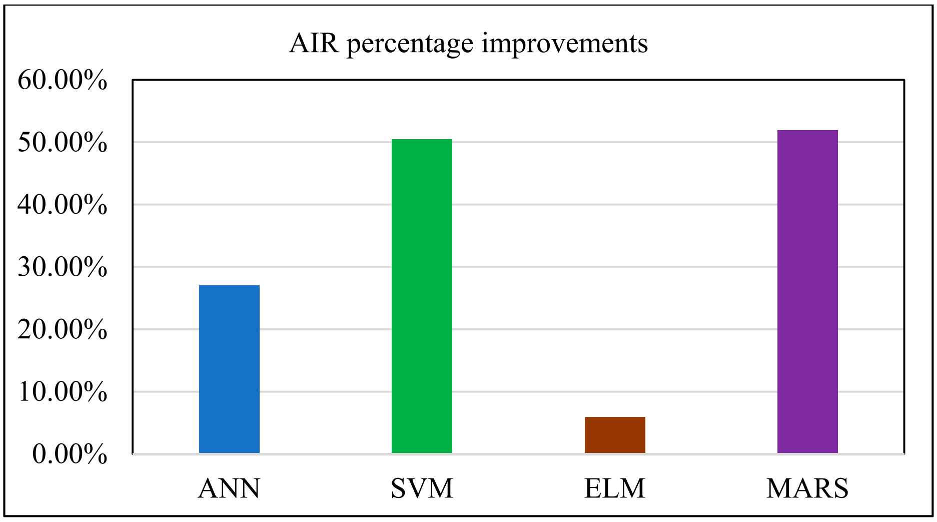

Figure 5. AIR percentage improvements for an MIMO process with three outputs, using the second scheme instead of the first scheme.

5. Conclusions

The automation (i.e., EPC) and information technology of big data are two of the primary concepts of the Industry 4.0. Therefore, intelligent and automated production systems that can effectively monitor multivariate quality characteristics of a process (i.e., MIMO system) have attracted considerable attention. When an MSPC signal is triggered, remedial actions that include the recognition of CCPs, determination of faults, and removal of the root causes should be taken to stabilize the process and return to in control conditions.

This paper is concerned with the recognition of mixture CCPs for MIMO systems. This task is particularly challenging since it encounters the difficulties of embedded and MCCPs in an MIMO system. This study presents several ML techniques for recognizing six and four types of MCCPs for two and three process outputs, respectively. Two schemes exist for designing the ML techniques. While there is no EPC involved in the first scheme, the second scheme includes EPC actions for all the ML classifiers. The performance of the proposed classification technique (i.e., the second scheme) is confirmed through a series of computer experiments. The proposed scheme maintains satisfactory performance in recognizing MCCPs for an MIMO process.

In this study, the MCCPs consist of four single CCPs, and an attempt to combine five or more single CCPs would be a valuable contribution of future research. However, one difficulty in classification design that will be encountered is that the number of categories of the classifiers' output nodes will increase when more types of MCCPs are involved in an MIMO system. Hybrid modeling techniques [31,32] and/or other ML techniques, such as artificial immune systems and random forests, may be worthy of implementation to decrease the number of output categories in the future.

Author Contributions: Conceptualization, Y.E.S.; Methodology, Y.E.S.; Software, Y.-T.H.; Writing and editing, Y.E.S. and Y.-T.H. All authors have read and agreed to the published version of the manuscript.

Funding: This work is partially supported by the Ministry of Science and Technology of the Republic of China (Taiwan), Grant No. 108-2221-E-030-005.

Acknowledgments: The authors would like to thank editors and anonymous reviewers for their careful reading and helpful remarks.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Shewhart, W.A. *Economic Control of Quality of Manufactured Product*; D. Van Nostrand Company, Inc.: New York, NY, USA, 1931.

2. Western Electric Company. *Statistical Quality Control Handbook*; Western Electric Company: New York, NY, USA, 1956.
3. Lu, C.J.; Shao, Y.E.; Li, P.H. Mixture control chart patterns recognition using independent component analysis and support vector machine. *Neurocomputing* **2011**, *74*, 1914–2011. [CrossRef]
4. De la Torre Gutiérrez, H.; Pham, D.T. Identification of patterns in control charts for processes with statistically correlated noise. *Int. J. Prod. Res.* **2018**, *56*, 1504–1520. [CrossRef]
5. Gutierrez, H.D.; Pham, D.T. Estimation and generation of training patterns for control chart pattern recognition. *Comput. Ind. Eng.* **2016**, *95*, 82–2016.
6. Xanthopoulos, P.; Razzaghi, T. A weighted support vector machine method for control chart pattern recognition. *Comput. Ind. Eng.* **2014**, *70*, 134–149. [CrossRef]
7. El-Midany, T.T.; El-Baz, M.A.; Abd-Elwahed, M.S. A proposed framework for control chart pattern recognition in multivariate process using artificial neural networks. *Expert Syst. Appl.* **2010**, *37*, 1035–1042. [CrossRef]
8. Ghomi, S.F.; Lesany, S.A.; Koochakzadeh, A. Recognition of unnatural patterns in process control charts through combining two types of neural networks. *Appl. Soft Comput.* **2011**, *11*, 5444–5456. [CrossRef]
9. Gauri, S.; Chakraborty, S. Improved recognition of control chart patterns using artificial neural networks. *Int. J. Adv. Manuf. Technol.* **2008**, *36*, 1191–1201. [CrossRef]
10. Ebrahimzadeh, A.; Addeh, J.; Rahmani, Z. Control chart pattern recognition using K-MICA clustering and neural networks. *ISA Trans.* **2012**, *51*, 111–119. [CrossRef]
11. Addeh, A.; Khormali, A.; Golilarz, N.A. Control chart pattern recognition using RBF neural network with new training algorithm and practical features. *ISA Trans.* **2018**, *79*, 202–216. [CrossRef]
12. Ebrahimzadeh, A.; Ranaee, V. Control chart pattern recognition using an optimized neural network and efficient features. *ISA Trans.* **2010**, *49*, 387–393. [CrossRef]
13. Zhou, X.; Jiang, P.; Wang, X. Recognition of control chart patterns using fuzzy SVM with a hybrid kernel function. *J. Intell. Manuf.* **2015**, *29*, 51–67. [CrossRef]
14. Lin, S.Y.; Guh, R.S.; Shiue, Y.R. Effective recognition of control chart patterns in autocorrelated data using a support vector machine based approach. *Comput. Ind. Eng.* **2011**, *61*, 1123–1134. [CrossRef]
15. Du, S.; Huang, D.; Lv, J. Recognition of concurrent control chart patterns using wavelet transform decomposition and multiclass support vector machines. *Comput. Ind. Eng.* **2013**, *66*, 683–695. [CrossRef]
16. Khormali, A.; Addeh, J. A novel approach for recognition of control chart patterns: Type-2 fuzzy clustering optimized support vector machine. *ISA Trans.* **2016**, *63*, 256–264. [CrossRef]
17. Yang, W.A.; Zhou, W.; Liao, W.; Guo, Y. Identification and quantification of concurrent control chart patterns using extreme-point symmetric mode decomposition and extreme learning machines. *Neurocomputing* **2015**, *147*, 260–270. [CrossRef]
18. Shao, Y.E.; Chiu, C.C. Applying emerging soft computing approaches to control chart pattern recognition for an SPC–EPC process. *Neurocomputing* **2016**, *201*, 28–2016. [CrossRef]
19. Shao, Y.E.; Chang, P.Y.; Lu, C.J. Applying two-stage neural network based classifiers to the identification of mixture control chart patterns for an SPC-EPC process. *Complexity* **2017**, *2017*, 10. [CrossRef]
20. Shao, Y.E.; Lin, S.C. Using a time delay neural network approach to diagnose the out-of-control signals for a multivariate normal process with variance shifts. *Mathematics* **2019**, *7*, 959. [CrossRef]
21. Alshraideh, H.; Khatatbeh, E. A gaussian process control chart for monitoring autocorrelated process data. *J. Qual. Technol.* **2014**, *46*, 317–322. [CrossRef]
22. Qiu, P.; Li, W.; Li, J. A new process control chart for monitoring short-range serially correlated data. *Technometrics* **2019**. [CrossRef]
23. Kadri, F.; Harrou, F.; Chaabane, S.; Sun, Y.; Tahon, C. Seasonal ARMA-based SPC charts for anomaly detection: Application to emergency department systems. *Neurocomputing* **2016**, *173*, 2102–2114. [CrossRef]
24. He, Q.P.; Wang, J. Statistical process monitoring as a big data analytics tool for smart manufacturing. *J. Process Control* **2017**, *67*, 35–43. [CrossRef]
25. Capaci, F.; Vanhatalo, E.; Kulahci, M.; Bergquist, B. The revised Tennessee Eastman process simulator as testbed for SPC and DoE methods. *Qual. Eng.* **2019**, *31*, 212–229. [CrossRef]
26. John, B.; Singhal, S. An application of integrated EPC–SPC methodology for simultaneously monitoring multiple output characteristics. *Int. J. Qual. Reliab. Manag.* **2019**, *36*, 669–685. [CrossRef]
27. Diao, G.; Zhao, L.; Yao, Y. A dynamic quality control approach by improving dominant factors based on improved principal component analysis. *Int. J. Prod. Res.* **2015**, *53*, 4287–4303. [CrossRef]

28. Yang, W.A.; Zhou, W. Autoregressive coefficient-invariant control chart pattern recognition in autocorrelated manufacturing processes using neural network ensemble. *J. Intell. Manuf.* **2015**, *26*, 1180–2015. [CrossRef]
29. Luo, M.; Zhang, K. A hybrid approach combining extreme learning machine and sparse representation for image classification. *Eng. Appl. Artif. Intell.* **2014**, *27*, 228–235. [CrossRef]
30. Cao, J.; Hao, J.; Lai, X.; Vong, C.M.; Luo, M. Ensemble extreme learning machine and sparse representation classification. *J. Frankl. Inst.* **2016**, *353*, 4526–4541. [CrossRef]
31. Shao, Y.E.; Hou, C.-D.; Chiu, C.-C. Hybrid intelligent modeling schemes for heart disease classification. *Appl. Soft Comput.* **2014**, *14*, 52–2014. [CrossRef]
32. Shao, Y.E.; Hou, C.-D. Change point determination for a multivariate process using a two-stage hybrid scheme. *Appl. Soft Comput.* **2013**, *13*, 1527–2013. [CrossRef]
33. Tseng, T.S.; Chou, J.R.; Lee, P.S. A study on a multivariate EWMA controller. *IIE Trans.* **2002**, *34*, 541–549. [CrossRef]
34. Tseng, S.T.; Mi, H.C.; Lee, I.C. A multivariate EWMA controller for linear dynamic processes. *Technometrics* **2016**, *58*, 104–115. [CrossRef]
35. Yang, L.; Sheu, S.H. Economic design of the integrated multivariate EPC and multivariate SPC charts. *Qual. Reliab. Eng. Int.* **2007**, *23*, 218–2007. [CrossRef]
36. Gauri, S.K.; Chakraborty, S. Feature-based recognition of control chart patterns. *Comput. Ind. Eng.* **2006**, *51*, 726–742. [CrossRef]
37. Vapnik, V.N. *The Nature of Statistical Learning Theory*; Springer: Berlin/Heidelberg, Germany, 2000.
38. Cherkassky, V.; Ma, Y. Practical selection of SVM parameters and noise estimation for SVM regression. *Neural Netw.* **2004**, *17*, 126–2004. [CrossRef]
39. Friedman, J.H. Multivariate adaptive regression splines (with discussion). *Ann. Stat.* **1991**, *19*, 141–1991.
40. Huang, G.B.; Zhu, Q.Y.; Siew, C.K. Extreme learning machine: Theory and applications. *Neurocomputing* **2006**, *70*, 501–2006. [CrossRef]
41. RStudio. Available online: https://rstudio.com/products/rstudio/ (accessed on 23 May 2017).
42. Hsu, C.W.; Chang, C.C.; Lin, C.J. *A Practical Guide to Support Vector Classification*; Technical Report; Department of Computer Science and Information Engineering, National Taiwan University: Taipei, Taiwan, 2003.