Hindawi
Complexity
Volume 2017, Article ID 2323082, 10 pages
https://doi.org/10.1155/2017/2323082

Research Article

Applying Two-Stage Neural Network Based Classifiers to the Identification of Mixture Control Chart Patterns for an SPC-EPC Process

Yuehjen E. Shao,[1] Po-Yu Chang,[1] and Chi-Jie Lu[2]

[1]*Department of Statistics and Information Science, Fu Jen Catholic University, New Taipei City, Taiwan*
[2]*Department of Industrial Management, Chien Hsin University of Science and Technology, Zhongli, Taoyuan County 32097, Taiwan*

Correspondence should be addressed to Yuehjen E. Shao; stat1003@mail.fju.edu.tw

Received 7 July 2017; Revised 6 September 2017; Accepted 19 September 2017; Published 22 October 2017

Academic Editor: Yanan Li

The effective controlling and monitoring of an industrial process through the integration of statistical process control (SPC) and engineering process control (EPC) has been widely addressed in recent years. However, because the mixture types of disturbances are often embedded in underlying processes, mixture control chart patterns (MCCPs) are very difficult for an SPC-EPC process to identify. This can result in problems when attempting to determine the underlying root causes of process faults. Additionally, a large number of categories of disturbances may be present in a process, but typical single-stage classifiers have difficulty in identifying large numbers of categories of disturbances in an SPC-EPC process. Therefore, we propose a two-stage neural network (NN) based scheme to enhance the accurate identification rate (AIR) for MCCPs by performing dimension reduction on disturbance categories. The two-stage scheme includes a combination of a NN, support vector machine (SVM), and multivariate adaptive regression splines (MARS). Experimental results reveal that the proposed scheme achieves a satisfactory AIR for identifying MCCPs in an SPC-EPC system.

1. Introduction

1.1. Background. High-quality products are typically manufactured through stable and effective industrial processes. Stable and effective processes are often complex, meaning that they must be well controlled and monitored using various advanced techniques. Although it is not new to industry, statistical process control (SPC) is still a widely used quality technique throughout industries. However, difficulty may be encountered when typical SPC charts are used to monitor an autocorrelated process. When the process measurements are autocorrelated, the false alarms are increased, and these improper signals can result in a misinterpretation. The use of engineering process control (EPC) has been successfully proposed to overcome this difficulty [1–4]. Accordingly, the SPC-EPC system plays an important role in industry. From a quality and control engineering viewpoint, SPC and EPC are two very important techniques for maintaining the quality of an industrial process. The main function of SPC is to monitor a process by plotting an observed sample on SPC control charts [5]. An out-of-control signal will be triggered when the process is unstable. The primary function of EPC is to compensate for the effects of process disturbances by manipulating controllable variables [1, 2]. Although they have their own individual features and specialties, SPC and EPC are naturally linked due to their roles in monitoring and controlling industrial processes. As a result, the goal-oriented integration of these techniques has been studied extensively in recent decades [1–4].

In an integrated mechanism, the role of SPC is to trigger an out-of-control signal when disturbances occur in a process and the role of EPC is to compensate for the effects of these disturbances. During this period, process personnel must identify the root causes of a disturbance and quickly fix them. However, the identification of root causes is difficult for an SPC-EPC process. The major reason for this is that although EPC can compensate for the effects of disturbances, it may conceal the mixture patterns of underlying disturbances.

Because the mixture control chart patterns (MCCPs) are hidden, it is much more difficult to identify the MCCPs in a complex system [6]. Additionally, because different mixture disturbances are typically associated with specific root causes that adversely affect the process, the effective identification of MCCPs for an SPC-EPC process is crucial [7]. Therefore, the issue of identification of MCCPs for industrial processes is an important research topic. Previous research has largely focused on the issue of identification of MCCPs for an SPC application alone. Therefore, the disturbances are assumed to be eliminated immediately after an SPC signal is triggered. In practical applications, the immediate elimination of disturbances is very unlikely. The identification of root causes of disturbances in industrial processes is complex and time-consuming. Consequently, this study mainly focuses on the identification of MCCPs when mixture disturbances exist in an SPC-EPC process.

1.2. Related Work. Due to its importance in the context of process improvement, many studies have investigated control chart pattern (CCP) recognition for industrial processes. For example, some have employed various NNs to recognize CCPs for processes [8, 9]. A NN ensemble-enabled model was used to classify unnatural CCPs for an autocorrelated process [10]. Process observations are assumed to follow an autoregressive with order 1 (i.e., AR(1)) process with an unknown constant coefficient. Because the autocorrelated process observations greatly affect the performance of NN based CCP classifiers, a learning vector quantization NN based approach was proposed to identify CCPs in an industrial process with various levels of autocorrelation [11]. Another team developed a two-module system for identifying six classes of CCPs [12]. The first module extracts features from the data and the second module contains several multilayer perceptron NNs with different parameters and training algorithms. Another group proposed a hybrid system that uses feature extraction modules, classifiers, and optimization modules to recognize common types of CCPs [13]. A multilayer perceptron NN and SVM were employed as classifiers.

In addition to using NN based classifiers to perform CCP recognition, SVM based classifiers are an effective alternative to identifying CCPs for a process. A multistage mechanism that combines independent component analysis and an SVM was proposed for classifying CCPs in a process in [14]. In addition to classifying CCPs for a process, one scheme was also reported to be capable of estimating abnormal pattern parameters by using an SVM and a probabilistic NN [15]. Because a typical SVM classifier may have poor performance, a weighted support vector machine (WSVM) approach was studied for classifying CCPs in an automated process in [16]. The research evaluated the performance of the WSVM approach on a real application from the wafer industry.

Several hybrid approaches have also been proposed for identifying CCPs in processes. A hybrid mechanism was used to combine wavelet filtering and adaptive resonance theory for the recognition of CCPs in a process in [17]. A hybrid approach combining extreme-point symmetric mode decomposition and an extreme learning machine (ELM) was studied in order to classify concurrent CCPs and accurately quantify major CCP parameters of the specific basic CCPs involved in [18]. A hybrid learning-based model was used to identify various types of unnatural CCPs in a process in [19]. A hybrid feature-based CCP classifier containing classification and regression trees, as well as NN techniques, was proposed for recognizing CCPs in [20].

In [21], it was reported that most related research has focused on single CCP recognition. Therefore, the research developed a hybrid methodology combining wavelet transforms and a particle swarm optimization-based support vector machine for concurrent CCP recognition. Because the assumption of a single type of unnatural pattern exists for the process data, poor CCP recognition may occur when unnatural concurrent CCPs are present in the process data. An approach combining singular spectrum analysis and a learning vector quantization network was proposed in order to recognize concurrent CCPs in [22]. Although this approach is useful for recognizing concurrent CCPs in univariate manufacturing processes, it did not account for the effects of using EPC on the process data. A thorough survey study regarding CCP recognition was performed in [23]. It was reported that a great number of CCP studies between 1991 and 2010 investigated various objectives and conditions. One can refer to [23] for additional details regarding related work on CCP recognition.

In spite of the considerable number of studies discussing the topic of CCPs, relatively little research has focused on methods of identifying a combination of two single CCPs or MCCPs in an SPC-EPC process. Although some have claimed that their approaches are suitable for CCP recognition in manufacturing process data [14, 22], they did not address the effects of EPC on the process data. In [6], although CCP recognition for an SPC-EPC system was investigated, the number of categories of CCPs was relatively small [6]. Additionally, the structure of the CCPs in [6] is of a consecutive type, meaning that the CCPs are easily recognized by one-stage classifiers. Therefore, this study aims to identify MCCPs in an SPC-EPC process. A first-order integrated moving average (IMA(1, 1)) component is assumed to be the noise in the process [1, 4]. Because minimum mean squared error (MMSE) has been successfully used in SPC-EPC processes, an MMSE is employed as the EPC component for the process [1, 4]. In this study, we arbitrarily assume that three individual disturbances can be present in the process simultaneously. Additionally, we focus on identifying the two single CCPs for a process. The reason for not considering single CCPs is that most studies have performed already identification tasks for such situations. The reason for not considering three single (i.e., whole single) CCPs is that we believe that the chance for having a whole single MCCP is very low.

A large number of categories (i.e., more than five) increase the degree of difficulty and lower the accurate identification rate (AIR) for the classification task. For example, in the case of five single disturbances in a process, the team in [6] reported that the values of AIR were 40.00%, 57.73%, 34.84%, and 54.40% with the use of NN, ELM, rough set, and random forest methods, respectively. In this study, because we assume that three single disturbances are present in a

process simultaneously, we have seven categories that must be classified. It is difficult to correctly classify seven categories using typical single-stage classifiers. Therefore, we propose a two-stage identification approach to overcome the problems associated with a large number of categories. Because a NN has data-driven and self-adaptive properties, as well as the ability to capture the nonlinear and complex underlying characteristics in an industrial process with a high degree of accuracy, we employ an NN classifier as the main component in the proposed two-stage mechanism. We also consider a support vector machine (SVM) and multivariate adaptive regression splines (MARS) as components for the proposed mechanism. The reason for choosing an SVM is that, other than ANNs, SVMs are one of the most widely used classifiers for CCP identification. The reason for choosing MARS is that they have not been adopted for CCP identification previously, although MARS are effective at classification and forecasting [24].

Experimental results reveal that the proposed two-stage NN based approach is able to effectively identify various MCCPs in an SPC-EPC process. The remainder of this paper is organized as follows. Section 2 discusses the structure of an industrial process and five types of single disturbances. The difficulty of MCCP identification in an SPC-EPC process is also addressed. The three soft computing techniques used for MCCP identification in this study are introduced in Section 3. Section 4 presents the results of simulation experiments to demonstrate the performance of the proposed approaches. The final section discusses the research findings and conclusions derived from this study.

2. The Process and Disturbance Models

2.1. The Industrial Process Models. A typical SPC process can be expressed as follows:

$$Y_t = u + a_t, \tag{1}$$

where Y_t is process output at time t and u is process mean level; without loss of generality, this study assumes that $u = 0$; a_t is white noise at time t; the white noise follows a normal probability distribution with mean of 0 and constant variance of σ^2. Without loss of generality, this study assumes that $\sigma^2 = 1$.

Because autocorrelation widely exists in practical chemical or continuous processes [10, 11, 25–27], the autocorrelation structure should be included in the process model. Therefore, the EPC is usually employed to compensate for effects of the autocorrelation and disturbances. A typical SPC-EPC process can be represented by a well-known zero-order process with the IMA(1, 1) noise [1, 4, 6]; that is,

$$\begin{aligned} Y_{t+1} &= qX_t + d_{t+1}, \\ d_{t+1} &= \frac{(1-\theta B)\, a_t}{(1-B)}, \end{aligned} \tag{2}$$

where d_{t+1} is process noise at time $t+1$, which follows an IMA(1, 1) process, θ is the parameter of an IMA(1, 1) process, X_t is control variable's measurement at time t, q is the system gain, which is a certain parameter, and B is backward shift operator, which is defined as $Y_t B^j = Y_{t-j}$ for $j = 1, 2, 3, \ldots$.

A suitable EPC or MMSE can be obtained when the process model is represented by (2), and it follows that [1]

$$X_t = -\frac{(1-\theta)}{q} \sum_{j=-\infty}^{t} Y_j. \tag{3}$$

2.2. The Disturbance Models. Disturbances may upset the process at any time. When a certain disturbance has occurred, the process model in (2) should be reformulated as

$$Y_{t+1} = qX_t + d_{t+1} + D_{t+1}, \tag{4}$$

where D_{t+1} is a certain disturbance at time $t+1$.

This study considers five single disturbances in an SPC-EPC process, and they are described as follows [28, 29]:

$$\begin{aligned} &\text{Cycle (CYC): } D_{t+1} = D^{\mathrm{CYC}}{}_t = \sin\left(\frac{2\pi t}{\psi}\right) U_t + a_t \\ &\text{Systematic (SYS): } D_{t+1} = D^{\mathrm{SYS}}{}_t = g \times (-1)^t + a_t \\ &\text{Shift (SHI): } D_{t+1} = D^{\mathrm{SHI}}{}_t = D_t + a_t \\ &\text{Stratification (STR): } D_{t+1} = D^{\mathrm{STA}}{}_t = r\sigma^2 \\ &\text{Trend (TRE): } D_{t+1} = D^{\mathrm{TRE}}{}_t = tS_t + a_t, \end{aligned} \tag{5}$$

where $D^{\mathrm{CYC}}{}_t$ is cycle disturbance value at time t, U_t is cycle amplitude, which is assumed to follow a uniform distribution within the range of (1.5, 2.5), ψ is cycle period, which is assumed to be $\psi = 8$. $D^{\mathrm{SYS}}{}_t$ is systematic disturbance value at time t, g is 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), $D^{\mathrm{SHI}}{}_t$ is shift disturbance value at time t, D_t is level of shift disturbance, which is assumed to be $D_t = 3$ after shifting, $D^{\mathrm{STA}}{}_t$ is stratification disturbance value at time t, r is random noise, which is assumed to follow a uniform distribution within the range of (0.2, 0.4), $D^{\mathrm{TRE}}{}_t$ is trend disturbance value at time t, and S_t is trend slope, which is assumed to follow a uniform distribution within the range of (0.05, 0.1).

3. The Problems and the Classifiers

3.1. The Problems. When one of those five disturbances is presented in the process, we refer to the presence of a single CCP. When any two of those five disturbances are concurrently presented in the process, we refer to the presence of MCCPs in this study. Figure 1 displays the process outputs when a single SHI disturbance has occurred after observation number 50 for an SPC process (i.e., (1)). Figure 2 shows the process outputs when a single SHI disturbance has occurred after observation number 50 for an SPC-EPC process (i.e., (2)). Because an MMSE or EPC is used for compensating for the disturbance's effects, we can observe that the pattern in Figure 2 is different from the pattern in

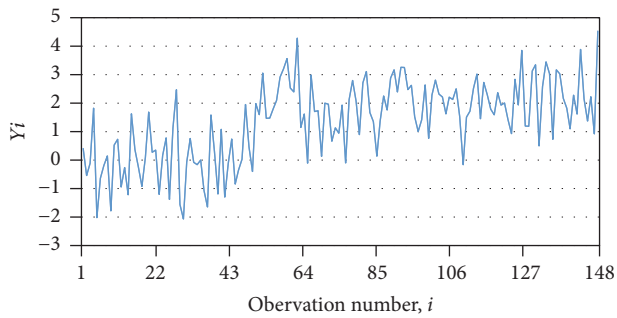

Figure 1: The SPC process outputs with the presence of a SHI disturbance intruding after observation number 50.

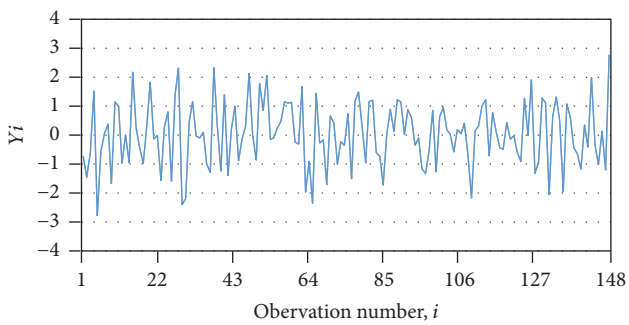

Figure 2: The SPC-EPC process outputs with the presence of a SHI disturbance intruding after observation number 50, using the EPC adjustments.

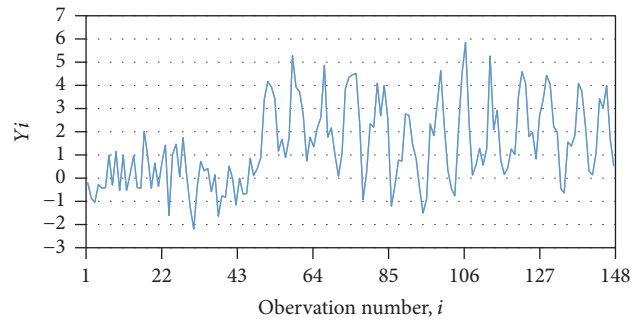

Figure 3: The SPC process outputs with the presence of a mixture type of SHI and CYC disturbances intruding after observation number 50.

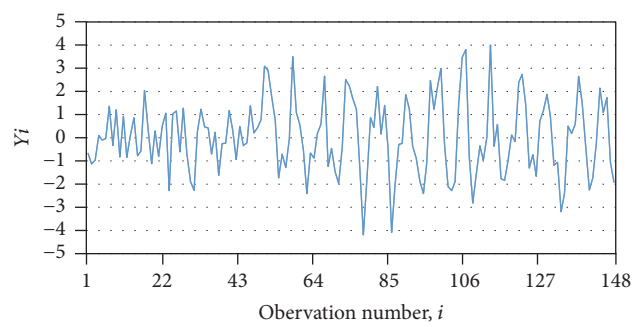

Figure 4: The SPC-EPC process outputs with the presence of a mixture type of SHI and CYC disturbances intruding after observation number 50, using the EPC adjustments.

Figure 1. Also, it is easier to identify the CCP in Figure 1, and most of the studies were concerned with the setting in Figure 1 (i.e., SPC alone). In contrast to the setting in Figure 1, this study focuses on the more complex setting in Figure 2 (i.e., an SPC-EPC process). The similar case would happen in other single disturbances as well.

When a single SHI disturbance and a single CYC disturbance concurrently occur, Figures 3 and 4 show the MCCPs for an SPC process and an SPC-EPC process, respectively. When a typical SPC chart is used to monitor the SPC-EPC system, an out-of-control signal will be triggered after observation number 51 in Figure 4. The process personnel can start searching for the root causes by investigating the CCPs. In addition, we can employ exponentially weighted moving average (EWMA) or cumulative sum (CUSUM) control charts to detect smaller magnitude of the disturbances [30, 31]. We can apply Shewhart type of control charts to the detection of a larger magnitude of the disturbances [5]. In addition, by comparison with CCPs in Figure 1 and Figure 3 or 4 for a process, we can notice that the MCCPs in Figure 3 or 4 is more complex to be recognized. By comparison with MCCPs in Figure 3 and Figure 4, the MCCPs in Figure 4 seem to be more difficult to be correctly recognized. Accordingly, this study aims to distinguish each individual single disturbance type for the MCCPs. One can obviously notice that the recognition of MCCPs in Figure 4 is a very difficult task.

Although many studies have investigated MCCPs identification, most of the existing works are concerned with the identification of the MCCPs for an SPC system alone. Only very few studies have focused on the CCP identification for a more sophisticated SPC-EPC process [6]. In [6], it was addressed that the five commonly observed disturbance patterns existed in an SPC-EPC process. The performances of the proposed extreme learning machine and random forest approaches were better than the approaches of ANN and rough set. In [6], the assumption of those five individual disturbances or their combined types of disturbances was consecutively occurred. That is, any two of the combined disturbances wound not intrude into the SPC-EPC system at the same time, and they just occurred in a sequential manner. For a generic concern, this study considers the case where the combined disturbances can occur at the same time. Also, since seven types of combined disturbance need to be considered in this study, the typical single-stage classifiers hardly perform the classification tasks well. Accordingly, this study proposes the two-stage NN based classifiers to overcome the above-mentioned difficulties.

3.2. NN Classifier. NN can be referred to as one of the most widely used classifiers for practical applications [32]. Because the backpropagation neural network (BPNN) is widely used in many applications, this study employs a BPNN when

designing the ANN model. NN modeling can be briefly described as follows. The relationship between output (y) and inputs $(x_1, x_2, \ldots, x_a)$ in an ANN model can be formed as

$$y = \alpha_0 + \sum_{j=1}^{b} \alpha_j g\left(\delta_{0j} + \sum_{i=1}^{a} \delta_{ij} x_i\right) + \varepsilon, \tag{6}$$

where α_j $(j = 0, 1, 2, \ldots, b)$ and δ_{ij} $(i = 0, 1, 2, \ldots, a;\ j = 0, 1, 2, \ldots, b)$ are the model connection weights; a is the number of input nodes; b is the number of hidden nodes; and ε is the error term.

In addition, a nonlinear functional mapping from the inputs $(x_1, x_2, \ldots, x_a)$ to the output y is performed by

$$y = f(x_1, x_2, \ldots, x_a, w) + \varepsilon, \tag{7}$$

where w is a vector of the model parameters and f is a function determined by the NN structure and connection weights.

For NN structure, this study employs a logistic function to serve as the transfer function in the hidden layer, and the logical function is represented as

$$g(z) = \frac{1}{1 + \exp(-z)}. \tag{8}$$

3.3. SVM Classifier. In addition to using NN classifiers, SVM played an important role for CCPs recognition for a process. SVM modeling can be described as follows. For SVM modeling, there are two separable classes and sample data can be described as

$$\begin{gathered} D = \{(x_1, y_1), \ldots, (x_l, y_l)\} \\ x_i \in \Re^n,\ y_i \in \{-1, 1\},\ i = 1, 2, \ldots, l \\ y_i = \begin{cases} 1, & \text{if } x_i \text{ in class } 1 \\ -1, & \text{if } x_i \text{ in class } 2, \end{cases} \end{gathered} \tag{9}$$

where l is the number of observations and n is the dimension of each observation. The decision function is given by $f(x) = \text{sign}(w^T x + b)$. The separating hyperplane is expressed as

$$f(x) = w^T x + b = 0, \tag{10}$$

where w is the coefficient vector and b is the constant.

To obtain the optimal hyperplane, we define the optimization problem as [33]

$$\begin{aligned} \min \quad & \Phi(\widetilde{x}) = \frac{1}{2} \|\widetilde{w}\|^2 \\ \text{s.t.} \quad & y_i\left(\widetilde{w}^T \widetilde{x}_i + b\right) \geq 1, \quad i = 1, 2, \ldots, n. \end{aligned} \tag{11}$$

It is difficult to solve (11), and we need to transform the optimization problem to the dual problem by Lagrange method. The value of α in the Lagrange method must be nonnegative real coefficients. Equation (11) is transformed into the following constrained form [33]:

$$\begin{aligned} \max \quad & \Phi(\widetilde{w}, b, \xi, \alpha, \beta) = \sum_{i=1}^{N} \alpha_i - \frac{1}{2} \sum_{i=1,j=1}^{N} \alpha_i \alpha_j y_i y_j \widetilde{x}_i^T \widetilde{x}_j \\ \text{s.t.} \quad & \sum_{j=1}^{n} \alpha_j y_j = 0, \quad 0 \leq \alpha_i \leq C,\ i = 1, 2, \ldots, n. \end{aligned} \tag{12}$$

In (12), C is the penalty factor and determines the degree of penalty assigned to an error.

In general, it could not find the linear separate hyperplane for all application data. For problems that cannot be linearly separated in the input space, the SVM uses the kernel method to transform the original input space into a high-dimensional feature space, where an optimal linear separating hyperplane can be found. The common kernel function are linear, polynomial, and radial basis function (RBF) and sigmoid. Since the RBF is one of the most widely used kernel functions, a RBF was used in this study. RBF is defined as [34]

$$K(\widetilde{x}_i, \widetilde{x}_j) = \exp\left(-\gamma \left\|\widetilde{x}_i - \widetilde{x}_j\right\|^2\right), \quad \gamma \geq 0, \tag{13}$$

where γ denotes the width of the RBF.

3.4. MARS Classifier. The MARS modeling is based on a divide-and-conquer strategy, where training datasets are partitioned into separate regions, each of which is assigned its own regression equation. The general MARS model can be described as follows [24]:

$$f(x) = b_0 + \sum_{m=1}^{M} b_m \prod_{j=1}^{J_m} \left[S_{jm}\left(x_{\nu(j,m)} - l_{jm}\right)\right], \tag{14}$$

where b_0 and b_m are the parameters, M is the number of basis functions (BFs), J_m is the number of knots, S_{jm} takes on values of either 1 or -1 and indicates the right or left sense of the associated step function, $\nu(j, m)$ is the label of the independent variable, and l_{jm} is the knot location. The optimal MARS model is selected in a two-step procedure. The first step is to build a large number of BFs to fit the data initially. The BFs are deleted in the order of least contributions to the most, using the generalized cross-validation (GCV) criterion in the second step. The measure of variable importance is provided by observing the decrease in the calculated GCV values when a variable is removed from the model. The GCV is described as follows:

$$\text{GCV}(M) = \frac{(1/N) \sum_{i=1}^{N} \left[y_i - f_M(x_i)\right]^2}{\left[1 - C(M)/N\right]^2}, \tag{15}$$

where N is the number of observations and $C(M)$ are the cost penalty measures of a model containing M BF.

3.5. Research Framework. Figure 5 displays a generalized depiction of the research framework. As shown in Figure 5, we notice that considerable process disturbances could

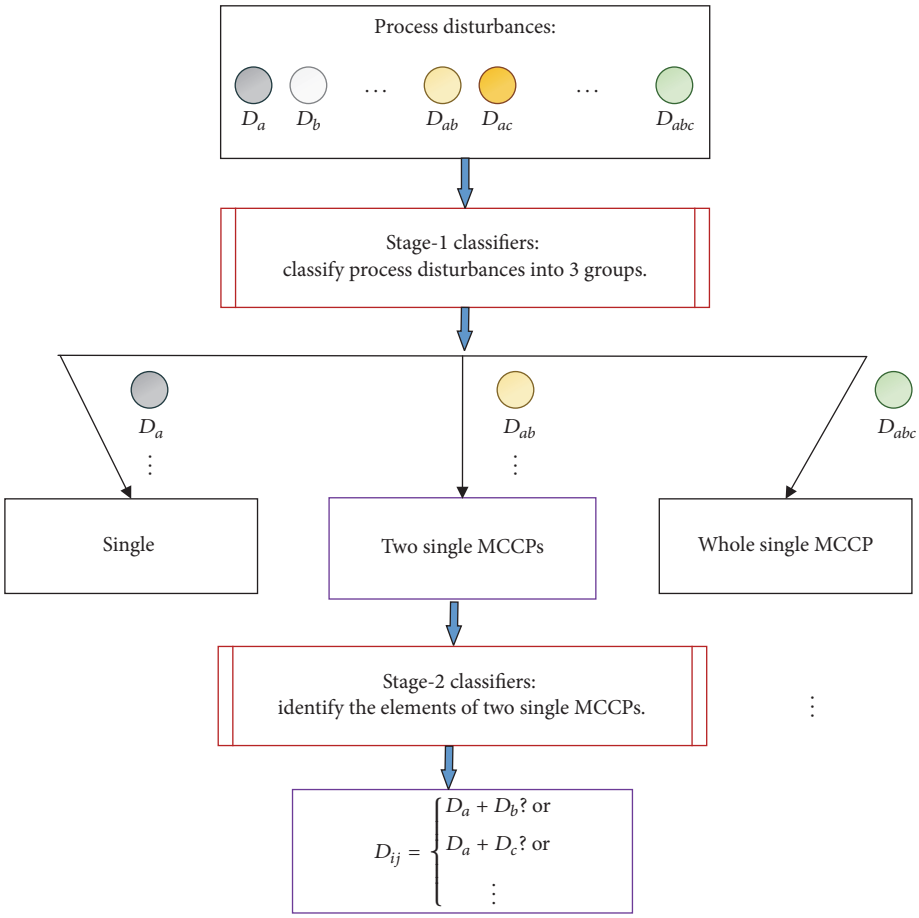

Figure 5: The research framework.

intrude into an SPC-EPC system. When disturbances intrude in the process, the process becomes unstable. In order to determine the root causes of the unstable process, we need to identify the patterns for the underlying disturbances. However, a classifier typically cannot perform the classification tasks well if the output variable has a large number of categories. Therefore, in stage-1, this study reduces the dimension of the output categories to three categories. These three categories of process disturbances include a single CCP, two single MCCPs, and whole single MCCP. Accordingly, the classifiers in stage-1 only need to classify fewer (i.e., three) categories of an output variable. Since the output variable of a classifier has fewer categories, the accuracy of classification can be greatly enhanced by using the classifiers in stage-1.

In addition, since this study focuses on identifying the two single MCCPs for the SPC-EPC process, the classifiers in stage-2 are used to identify elements (i.e., single disturbances) consisting of the two single MCCPs. Because fewer categories are associated with the output variables for stage-2 classifiers, the classifiers can have greater chance to maintain the high classification accuracy. Also, in this study, we employ three classifiers, ANN, SVM, and MARS, to perform the classification tasks in stage-1 and stage-2.

4. Experimental Results and Discussion

In this study, an industrial SPC-EPC process is assumed to be disturbed by five single disturbances that are described by (5). Also, since this study assumes that any 3 single disturbances may concurrently be intruded into the process, there are 7 kinds or categories that may need to be classified. For example, suppose that there are three single disturbances, CYC, SYS, and SHI, concurrently intruding into a process; Table 1 shows three types with seven categories of CCPs. Here, this study is mainly interested in identifying the two single mixture MCCPs for the SPC-EPC process. That is, this study focuses on identifying three MCCPs: {CYC, SYS}, {CYC, SHI}, and {SYS, SHI}.

In this study, the NN simulator Qnet97, which was developed by Vesta Services Inc., was used to develop the NN models. Qnet97 is a C-based simulator that provides a system for developing BPNN configurations by using a generalized delta learning algorithm. For SVM modeling, the package "e1071" in the R programming language was used in this study. The MARS model was constructed using MARS, which was developed by Salford Systems.

In addition, since seven categories are difficult to be correctly identified by using typical single-stage classifiers,

Table 1: Types of CCPs for 3 single disturbances in a process.

Types	Categories	Combination
Type I		
Single	C_1	{CYC}
Single	C_2	{SYS}
Single	C_3	{SHI}
Type II		
Two single MCCPs	C_4	{CYC, SYS}
Two single MCCPs	C_5	{CYC, SHI}
Two single MCCPs	C_6	{SYS, SHI}
Type III		
Whole single MCCP	C_7	{CYC, SYS, SHI}

Table 2: The confusion matrix for the ANN classifier.

Predicted class	Actual class						
	0	1	2	3	4	5	6
0	**20**	6	0	0	0	0	10
1	200	**47**	0	0	0	0	0
2	36	223	**55**	70	43	0	55
3	5	24	245	**186**	93	0	100
4	6	0	0	44	**164**	0	120
5	33	0	0	0	0	**70**	15
6	0	0	0	0	0	230	**0**

this study proposes a two-stage identification mechanism to overcome the problem of considerable categories. The first stage of the proposed two-stage classifiers is used to initially identify three types of disturbances, type I, type II, and type III. Instead of identifying seven categories of disturbances, the classifiers in the first stage only need to identify three types of disturbances with the use of dimension reduction approach. Then, the second stage of the proposed classifiers is served to identify three categories of disturbances in type II. Since this study aims to obtain the AIR values for the three categories, C_4, C_5, and C_6, we need to know the calculation of these AIR values. In this study, AIR is used for the classification performance measurement. AIR is defined as follows:

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

where N is the total number of data vectors used in the identification process and n_a is the number of data vectors in N, where the true CCP type is accurately identified.

Considering the case of C_4, we can have the AIR by using the following procedure. By using a certain classifier of the proposed mechanism, we can obtain the AIR values, denoted as A_1 and A_2, in the first and second stages, respectively. The AIR for the case of C_4 is simply obtained by the multiplication of A_1 and A_2. The AIR values for C_5 and C_6 can be easily obtained by using the same procedures.

In order to demonstrate the identification capability of the proposed approaches, this study performs a series of computer simulations. Suppose that an SPC-EPC process is represented by (4) and the parameters are arbitrarily selected as $q = 0.5$ and $\theta = 0.8$. In this study, we consider that three out of five single disturbances will intrude into an SPC-EPC system. Accordingly, we have seven combinations of CCPs for each set of concurrent three single disturbances. For example, by observing Table 1, we can have seven combinations for the concurrent three single disturbances. Based on the SPC-EPC system and disturbance models, this study generates the data vectors for the training and testing phases of the NN, SVM, and MARS classifiers. For the structures of these three classifiers, this study employs X (i.e., (3)) and Y (i.e., (4)) as the classifiers' inputs and considers Z (i.e., the classification category) as the classifiers' output. Since the typical classifiers would identify seven combinations, the values of Z are set to be from 0 to 6. The meanings of these values are described as follows:

$Z = 0$ represents the presence of the first combination (i.e., single {CYC}).

$Z = 1$ represents the presence of the second combination (i.e., single {SYS}).

$Z = 2$ represents the presence of the third combination (i.e., single {SHI}).

$Z = 3$ represents the presence of the fourth combination (i.e., MCCP {CYC-SYS}).

$Z = 4$ represents the presence of the fifth combination (i.e., MCCP {CYC-SHI}).

$Z = 5$ represents the presence of the sixth combination (i.e., MCCP {SYS-SHI}).

$Z = 6$ represents the presence of the seventh combination (i.e., whole single MCCP {CYC-SYS-SHI}).

Additionally, this study uses 4,900 and 2,100 data vectors for the training and testing phases, respectively. In the training phase, a set of 700 data vectors are generated from each combination. Consider Table 1 as an example. The first 700 data vectors are generated from the presence of {CYC} alone. The data vectors from 701 through 1,400 are generated from the presence of {SYS} alone. The same grouping is used up to the final data vectors from 4,201 through 4,900, which are generated from the combined presence of {CYC, SYS, SHI} disturbances. The testing data structure is similar to the training data structure. Specifically, the first 300 data vectors involve {CYC} disturbances alone and the final data vectors from 1,801 through 2,100 involve {CYC, SYS, SHI} disturbances. After performing the classification tasks with ANN, SVM, and MARS classifiers, we can compute the corresponding type I error and type II error rates. Tables 2, 3, and 4 present the corresponding confusion matrices for ANN, SVM, and MARS, respectively. For the identification results for all the categories in Table 1, type I error rates are 0.7419, 0.5767, and 0.4933 for ANN, SVM, and MARS, respectively. Also, type II error rates are 0.1237, 0.0961, and 0.0822 for ANN, SVM, and MARS, respectively. We can notice that all type I and type II error rates are not satisfactory.

In comparison to type I and type II errors, the AIR measure is easier to be understood by process personnel.

Table 3: The confusion matrix for the SVM classifier.

Predicted class	Actual class						
	0	1	2	3	4	5	6
0	**151**	201	1	64	44	1	77
1	137	**99**	0	5	0	0	0
2	1	0	**260**	15	33	0	36
3	2	0	0	**88**	17	5	2
4	5	0	25	54	**156**	8	0
5	3	0	0	43	12	**272**	0
6	1	0	14	31	38	14	**185**

Table 4: The confusion matrix for the MARS classifier.

Predicted class	Actual class						
	0	1	2	3	4	5	6
0	**90**	150	0	41	38	0	0
1	193	**149**	0	5	0	0	0
2	5	0	**293**	29	62	0	60
3	1	1	0	**70**	1	0	105
4	7	0	7	61	**164**	2	52
5	4	0	0	75	30	**298**	83
6	0	0	0	19	5	0	**0**

Additionally, since AIR was employed in [6], this study uses the AIR as a measure of accuracy for the various two-stage NN based classifiers presented in our study. Table 5 presents the identification results for all the CCPs in Table 1. The overall AIR values are 25.81%, 57.67%, and 50.76% for the ANN, SVM, and MARS classifiers, respectively. The AIR values of the three MCCPs (i.e., {CYC, SYS}, {CYC, SHI}, and {SYS, SHI}) are 46.47%, 57.33%, and 59.11% for the ANN, SVM, and MARS classifiers, respectively. These AIR values can be computed by using a confusion matrix. The overall AIR for the ANN model is computed by summing the diagonal elements and dividing the testing data vectors (i.e., 2,100) as follows:

$$\text{overall AIR} = \frac{\sum (20 + 47 + \cdots + 0)}{300 \times 7} = 25.81\%. \tag{17}$$

Additionally, because the class values of 3, 4, and 5 in Table 2 represent the status of three MCCPs (i.e., {CYC, SYS}, {CYC, SHI}, and {SYS, SHI}), the AIR of the ANN model for the three MCCPs is computed as follows:

$$\text{AIR} = \frac{\sum (186 + 164 + 70)}{300 \times 3} = 46.67\%. \tag{18}$$

By using the same calculations, we can obtain the corresponding AIR values for the SVM and MARS models.

After performing ANN modeling, we found that a {2-6-1} topology with a learning rate of 0.01 provided the best results with the minimum testing RMSE. The notation $\{n_i\text{-}n_h\text{-}n_o\}$ represents the number of neurons in the input layer, hidden layer, and output layer, respectively. Because the RBF kernel function is adopted in this study, the performance of the SVM is mainly affected by the values of two parameters (C and γ). There are no general rules for the choice of C and γ. In this study, the grid search proposed in [35] is used for parameter settings. The grid search method uses exponentially growing sequences of C and γ to identify good parameters (e.g., $C = 2^{-5}, 2^{-3}, 2^{-1}, \ldots, 2^{15}$). The parameter settings for C and γ that generate the highest correct classification rate are considered to be ideal set. The trained SVM model with the best parameter settings, denoted as $\{C, \gamma\}$, is preserved and used during the monitoring stage for CCP recognition. Additionally, because there are no specific parameter settings for MARS, we simply denote the parameter settings as {null} for the MARS classifiers.

In a traditional design, the identification performance of the three classifiers is poor due to the fact that the classifier output Z contains too many categories. Therefore, we propose a two-stage mechanism in order to overcome the problems associated with a large number of categories.

Since identification performance of the typical design is not satisfactory when the output categories are considerably large, this study reduces the dimension of the output categories by using a two-stage mechanism. By considering Table 1, the first stage of the classifiers is used to identify three, instead of seven, types of disturbances (i.e., type I, type II, and type III). In the first stage, this study initially sets the values of classification categories Z as 0, 1, and 2, respectively. The value of 0 represents the presence of type I disturbances (i.e., single disturbance), the value of 1 represents the presence of type II disturbances (i.e., the MCCPs which we want to classify), and the value of 2 represents the presence of type III disturbances (i.e., the whole single CCP). This study also uses 4900 and 2100 data vectors for the training and testing phases. The second stage of the proposed design is to identify which set of MCCPs are presented in the underlying process. Namely, which one of {CYC, SYS}, {CYC, SHI}, and {SYS, SHI} existed in the system? In the second stage, since the process contains three combinations of MCCPs in type II, this study sets Z as three output values. "0" indicates that {CYC, SYS} is presented, "1" indicates that {CYC, SHI} is presented, and "2" indicates that the {SYS, SHI} mixture disturbance is presented in the process.

After performing the two-stage classification tasks, this study obtains the results which are listed in Table 6. In Table 6, the first column lists ten combinations for three out of five single disturbances. The second column presents the best two-stage classifiers associated with the parameter settings for the first and second stages, respectively. The last column in Table 6 shows the average AIR values. Observing Table 6, we can notice that the smallest average AIR value is 57.23%, occurring in the case of {SHI, TRE, CYC}. The possible reason may be the fact that the characteristics for the SHI, TRE, and CYC are similar, and those three disturbances cannot be effectively identified. In addition, all other average AIR values for the remaining nine combinations are greater than 60%. In general, the proposed BPNN-SVM and BPNN-BPNN classifiers possess satisfactory capability for identifying the MCCPs for an SPC-EPC process.

Table 5: Identification results for a single-stage design when {CYC, SYS, SHI} exist concurrently in an SPC-EPC process.

Classifier {parameter settings}	Overall AIR	AIR for three MCCPs
ANN {2, 6, 1}	25.81%	46.67%
SVM $\{2^{-2}, 2^{5}\}$	57.67%	57.33%
MARS {null}	50.67%	59.11%

Table 6: Identification results for all combinations of MCCPs in a two-stage design.

Combinations	Best two-stage classifiers [{parameters}, {parameters}]	Average AIR
{CYC, SYS, STR}	BPNN-SVM [{2, 2, 1}, {0.0625, 0.5}]	73.75%
{SHI, SYS, CYC}	SVM-SVM [{0.0625, 8}, {0.5, 8}]	62.89%
{SHI, SYS, STR}	BPNN -SVM [{2, 5, 1}, {2, 4}]	84.89%
{SHI, CYC, STR}	BPNN -MARS [{2, 6, 1}, {null}]	80.69%
{TRE, SYS, CYC}	BPNN -SVM [{2, 5, 1}, {0.5, 32}]	60.58%
{TRE, SYS, STR}	BPNN -MARS [{2, 5, 1}, {null}]	73.89%
{TRE, CYC, STR}	BPNN - BPNN [{2, 3, 1}, {2, 6, 1}]	77.72%
{SHI, TRE, CYC}	BPNN -SVM [{2, 4, 1}, {1, 8}]	57.23%
{SHI, TRE, SYS}	BPNN - BPNN [{2, 5, 1}, {2, 6, 1}]	79.22%
{SHI, TRE, STR}	BPNN - BPNN [{2, 3, 1}, {2, 6, 1}]	61.09%

5. Conclusion

CCP identification is crucial for the improvement of industrial processes. Because an integrated mechanism using SPC and EPC can result in very effective monitoring and controlling performance, we must focus heavily on CCP identification for such a process. Thus, the purpose of this study was to identify MCCPs for an SPC-EPC system. Additionally, we proposed a two-stage classification technique in order to overcome the problems associated with a large number of output categories. The first stage of the proposed approach employs classifiers to identify a reduced number of output categories and the second-stage classifiers are used to effectively determine the types of MCCPs for a process.

The performance of the proposed two-stage classification technique was verified through a series of computer experiments. The proposed BPNN-SVM and BPNN-BPNN models achieve satisfactory performance for identifying the MCCPs for an SPC-EPC process. In our study, we used the AIR to measure accuracy for various two-stage NN based classifiers. Another measurement, area under the receiver operating characteristic (ROC) curve (AUC), could also be used to measure the accuracy of various classifiers. One limitation of such a measure is that AUC values are difficult to calculate for our proposed two-stage models. One possible future research direction is to compute AUC values for our two-stage models. Another limitation for the proposed two-stage classifiers is the computational time. A two-stage classifier may require more time to obtain the classification results. Faster computer systems are suggested to perform the two-stage classification tasks, since they would help to speed up the process.

Additionally, this study considered two single MCCPs identification. An attempt to classify three single or even four single MCCPs would be a valuable contribution to future research. Some other classifiers, such as multivariate adaptive regression splines (MARS) and random forests may also be employed to identify the mixture disturbance patterns for a multivariate SPC-EPC system.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this article.

Acknowledgments

This work is partially supported by the Ministry of Science and Technology of the Republic of China (Grant no. MOST 106-2221-E-030-010-MY2).

References

[1] J. F. MacGregor, T. J. Harris, and J. D. Wright, "Duality between the control of processes subject to randomly occurring deterministic disturbances and arima stochastic disturbances," *Technometrics*, vol. 26, no. 4, pp. 389–397, 1984.

[2] G. Box and T. Kramer, "Statistical process monitoring and feedback adjustment—a discussion," *Technometrics*, vol. 34, no. 3, pp. 251–285, 1992.

[3] D. C. Montgomery and C. M. Mastrangelo, "Some statistical process control for autocorrelation data (with discussion)," *Journal of Quality Technology*, vol. 23, no. 3, pp. 179–193, 1991.

[4] Y. E. Shao, "Integrated application of the cumulative score control chart and engineering process control," *Statistica Sinica*, vol. 8, no. 1, pp. 239–252, 1998.

[5] W. A. Shewhart, *Economic Control of Quality of Manufactured Product*, D. Van Nostrand Company, Inc., New York, NY, USA, 1931.

[6] Y. E. Shao and C.-C. Chiu, "Applying emerging soft computing approaches to control chart pattern recognition for an SPC–EPC process," *Neurocomputing*, vol. 201, pp. 19–28, 2016.

[7] C.-J. Lu, Y. E. Shao, and P.-H. Li, "Mixture control chart patterns recognition using independent component analysis and support vector machine," *Neurocomputing*, vol. 74, no. 11, pp. 1908–1914, 2011.

[8] R.-S. Guh and J. D. T. Tannock, "Recognition of control chart concurrent patterns using a neural network approach," *International Journal of Production Research*, vol. 37, no. 8, pp. 1743–1765, 1999.

[9] S. A. Lesany, A. Koochakzadeh, and S. M. T. Fatemi Ghomi, "Recognition and classification of single and concurrent unnatural patterns in control charts via neural networks and fitted line of samples," *International Journal of Production Research*, vol. 52, no. 6, pp. 1771–1786, 2014.

[10] W.-A. Yang and W. Zhou, "Autoregressive coefficient-invariant control chart pattern recognition in autocorrelated manufacturing processes using neural network ensemble," *Journal of Intelligent Manufacturing*, vol. 26, no. 6, pp. 1161–1180, 2015.

[11] R.-S. Guh, "Real-time recognition of control chart patterns in autocorrelated processes using a learning vector quantization network-based approach," *International Journal of Production Research*, vol. 46, no. 14, pp. 3959–3991, 2008.

[12] V. Ranaee and A. Ebrahimzadeh, "Control chart pattern recognition using neural networks and efficient features: a comparative study," *Pattern Analysis and Applications*, vol. 16, no. 3, pp. 321–332, 2013.

[13] A. Ebrahimzadeh, J. Addeh, and V. Ranaee, "Recognition of control chart patterns using an intelligent technique," *Applied Soft Computing*, vol. 13, no. 5, pp. 2970–2980, 2013.

[14] L.-J. Kao, T.-S. Lee, and C.-J. Lu, "A multi-stage control chart pattern recognition scheme based on independent component analysis and support vector machine," *Journal of Intelligent Manufacturing*, vol. 27, no. 3, pp. 653–664, 2016.

[15] H. De La Torre Gutierrez and D. T. Pham, "Estimation and generation of training patterns for control chart pattern recognition," *Computers & Industrial Engineering*, vol. 95, pp. 72–82, 2016.

[16] P. Xanthopoulos and T. Razzaghi, "A weighted support vector machine method for control chart pattern recognition," *Computers & Industrial Engineering*, vol. 70, no. 1, pp. 134–149, 2014.

[17] C.-H. Wang and W. Kuo, "Identification of control chart patterns using wavelet filtering and robust fuzzy clustering," *Journal of Intelligent Manufacturing*, vol. 18, no. 3, pp. 343–350, 2007.

[18] W.-A. Yang, W. Zhou, W. Liao, and Y. Guo, "Identification and quantification of concurrent control chart patterns using extreme-point symmetric mode decomposition and extreme learning machines," *Neurocomputing*, vol. 147, no. 1, pp. 260–270, 2015.

[19] M. Salehi, A. Bahreininejad, and I. Nakhai, "On-line analysis of out-of-control signals in multivariate manufacturing processes using a hybrid learning-based model," *Neurocomputing*, vol. 74, no. 12-13, pp. 2083–2095, 2011.

[20] S. K. Gauri and S. Charkaborty, "Recognition of control chart patterns using improved selection of features," *Computer & Industrial Engineering*, vol. 56, pp. 1577–1588, 2009.

[21] S. C. Du, D. L. Huang, and J. Lv, "Recognition of concurrent control chart patterns using wavelet transform decomposition and multiclass support vector machines," *Computers & Industrial Engineering*, vol. 66, no. 4, pp. 683–695, 2013.

[22] N. Gu, Z. Cao, L. Xie, D. Creighton, M. Tan, and S. Nahavandi, "Identification of concurrent control chart patterns with singular spectrum analysis and learning vector quantization," *Journal of Intelligent Manufacturing*, vol. 24, no. 6, pp. 1241–1252, 2013.

[23] W. Hachicha and A. Ghorbel, "A survey of control-chart pattern-recognition literature (1991–2010) based on a new conceptual classification scheme," *Computers & Industrial Engineering*, vol. 63, no. 1, pp. 204–222, 2012.

[24] J. H. Friedman, "Multivariate adaptive regression splines," *The Annals of Statistics*, vol. 19, no. 1, pp. 1–67, 1991.

[25] S. Psarakis, "The use of neural networks in statistical process control charts," *Quality and Reliability Engineering International*, vol. 27, no. 5, pp. 641–650, 2011.

[26] F. Kadri, F. Harrou, S. Chaabane, Y. Sun, and C. Tahon, "Seasonal ARMA-based SPC charts for anomaly detection: application to emergency department systems," *Neurocomputing*, vol. 173, pp. 2102–2114, 2016.

[27] M. Weese, W. Martinez, F. M. Megahed, and L. A. Jones-Farmer, "Statistical learning methods applied to process monitoring: an overview and perspective," *Journal of Quality Technology*, vol. 48, no. 1, pp. 4–27, 2016.

[28] S. K. Gauri and S. Chakraborty, "Feature-based recognition of control chart patterns," *Computers & Industrial Engineering*, vol. 51, no. 4, pp. 726–742, 2006.

[29] S. Y. Lin, R. S. Guh, and Y. R. Shiue, "Effective recognition of control chart patterns in autocorrelated data using a support vector machine based approach," *Computers & Industrial Engineering*, vol. 61, pp. 1123–1134, 2011.

[30] S. W. Roberts, "Control chart tests based on geometric moving averages," *Technometrics*, vol. 1, no. 3, pp. 239–250, 1959.

[31] E. S. Page, "Continuous inspection schemes," *Biometrika*, vol. 41, no. 1-2, pp. 100–115, 1954.

[32] Y. E. Shao, C.-D. Hou, and C.-C. Chiu, "Hybrid intelligent modeling schemes for heart disease classification," *Applied Soft Computing*, vol. 14, pp. 47–52, 2014.

[33] V. N. Vapnik, *The Nature of Statistical Learning Theory*, Springer, Berlin, Germany, 2000.

[34] V. Cherkassky and Y. Ma, "Practical selection of SVM parameters and noise estimation for SVM regression," *Neural Networks*, vol. 17, no. 1, pp. 113–126, 2004.

[35] C. W. Hsu, C. C. Chang, and C. J. Lin, "A practical guide to support vector classification," Tech. Rep., Department of Computer Science and Information Engineering, National Taiwan University, Taipei, Taiwan, 2003.