REGULAR PAPER

Cyber Physical Energy System for Saving Energy of the Dyeing Process with Industrial Internet of Things and Manufacturing Big Data

Kyu Tae Park[1] · Yong Tae Kang[2] · Suk Gon Yang[3] · Wen Bin Zhao[1] · Yong-Shin Kang[1] · Sung Ju Im[1] · Dong Hyun Kim[2] · Su Young Choi[2] · Sang Do Noh[1]

Abstract

The manufacturing industry has recently been focusing on improving energy efficiency to reduce greenhouse gas emissions and achieve sustainable growth. The focus is on combining existing energy technologies with new information and communication technologies as the Fourth Industrial Revolution approaches. Dyeing and finishing shops use the most energy in the textile industry and have below-average energy efficiency because of low technology and facility investment. Research into increasing the energy efficiency of dyeing and finishing shops has concentrated on developing equipment; however, it is difficult for small- and medium-sized factories to benefit from these advances. Thus, research into means of improving energy efficiency by improving the process and system efficiency of dyeing and finishing companies, who have difficulties with facility investment and operation, is necessary. In this study, a cyber physical energy system that improves the energy efficiency of the dyeing process by collecting and analyzing manufacturing big data was developed. Further, the implemented system was applied to actual dyeing and finishing shops, and its effects were verified. This research contributes to improving the situation of the dyeing process using machine learning techniques and manufacturing big data by adjusting cyber physical energy systems without utilizing expensive equipment. Inaccurate process instruction from the laboratory are also replaced by the cyber physical energy system, and the invalid and inefficient steps in the traditional process scenario derived from operator's experience are replaced with more valid and usable actions.

Keywords Big data · Cyber physical system · Dyeing process · Energy efficiency · Green manufacturing · Industrial internet of things

✉ Sang Do Noh
   sdnoh@skku.edu

   Kyu Tae Park
   parkkyutae0201@gmail.com

   Yong Tae Kang
   ytkang@coever.co.kr

   Suk Gon Yang
   sukgon@dyetec.or.kr

   Wen Bin Zhao
   zhaowenbin86@gmail.com

   Yong-Shin Kang
   yskang7867@gmail.com

   Sung Ju Im
   realdream@skku.edu

   Dong Hyun Kim
   kdh@coever.co.kr

   Su Young Choi
   csy@coever.co.kr

[1] Department of Industrial Engineering, SungKyunKwan University, 2066, Seobu-ro, Jangan-gu, Suwon-si, Gyeonggi-do 16419, Republic of Korea

[2] CoEver I & T Co., Ltd, 168, Gasan Digital 1-ro, Geumcheon-gu, Seoul 08507, Republic of Korea

[3] Korea Dyeing & Finishing Technology Institute, 92, Dalseocheon-ro, Seo-gu, Daegu 41706, Republic of Korea

1 Introduction

As part of efforts to realize green manufacturing, the manufacturing society has been focusing on developing technologies to improve energy efficiency and sustainability [1–4] Increasing energy efficiency has become a major consideration owing to rising global carbon dioxide emissions and energy resource prices and climate change. Companies are trying to strengthen their competitiveness through innovations in such technologies [1, 3, 5].

Energy-related research in the manufacturing industry has focused on developing new energy sources, purifying fossil fuels, and improving process and system efficiency. The first two directions have been the traditional focus, but studies on improving process and system efficiency have been attracting attention because they should also improve major indicators, such as the productivity of the whole system [1, 6].

Diverse and convergent technologies, such as industrial internet of things (IIoT) and big data application, have been developed and applied in accordance with manufacturing innovation paradigms, such as Industry 4.0 [6, 7]. As an example, the cyber physical energy system (CPES) has been applied to a wide range of industries because it increases energy efficiency through the collection, processing, and utilization of information and through the application of modeling techniques [7, 8]. The CPES is an eco-friendly manufacturing concept because it can reduce fuel use and environmental pollution by improving energy efficiency through energy use and production process optimization as well as process efficiency [7–9].

The textile industry has lower energy efficiency and needs much more energy efficiency improvement than other manufacturing industries. In the textile industry, dyeing and finishing shops have the highest energy consumption, accounting for 42% of the total energy consumption. It is also an industry that has a lot of energy-related expenses compared to the total cost. Thus, research has been focused on improving the energy efficiency of dyeing and finishing shops [10, 11]. However, such studies mainly focused on the development of new equipment; thus, they only benefit factories with sufficient capital and not the small- and medium-sized factories that are incapable of large-scale investment. In addition, because studies to improve the energy efficiency of dyeing and finishing shops through process efficiency enhancement are insufficient, research is needed on dyeing and finishing systems with high usability and advanced production methods to enhance the energy efficiency of small- and medium-sized companies [4, 11–13]. Deep understanding and sufficient consideration of the actual site are required for the effective application of such systems and the development of dard production and management methods [12, 14].
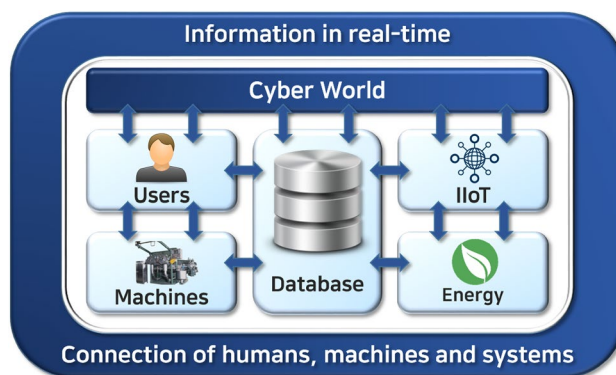

Fig. 1 Conceptual diagram of a cyber physical energy system

In this study, a method that improves the energy efficiency of the dyeing process was developed based on the CPES concept. The CPES of this study was implemented by assigning the research domain to the dyeing process, which is the main energy-consuming process in dyeing and finishing shops. The study was conducted with the aim of identifying and solving the process inefficiencies of the current dyeing process and reducing the possibility of repeated-dyeing, in which the dyeing process is performed again because of defects in the dyeing process. The CPES implemented in this study prevents improper use of energy by improving the process and system efficiency without the need to invest in expensive equipment.

The following tasks will be performed herein: (1) problems with increasing the energy efficiency of dyeing and finishing are defined; (2) solutions and scenarios for the problems are presented (Fig. 1); (3) the design of the CPES architecture based on these solutions and scenarios is explained; (4) the data model for manufacturing big data collected using IIoT devices and retrieved from databases is defined and characterized; and (5) modules designed with machine learning techniques using manufacturing big data are presented and (6) validated through a case study with actual dyeing and finishing shops.

2 Research Background

2.1 Cyber Physical Energy System

A cyber physical system (CPS) is a physical and engineered system that performs monitoring, controlling, and coordinating using computing technology and ICT. A CPS consists of a physical world, which is the actual site, and a cyber world, which is constructed using information and knowledge [14]. A CPS supports decision making related to the manufacturing process by analyzing and reflecting on the complex situation of the physical world through data collection,
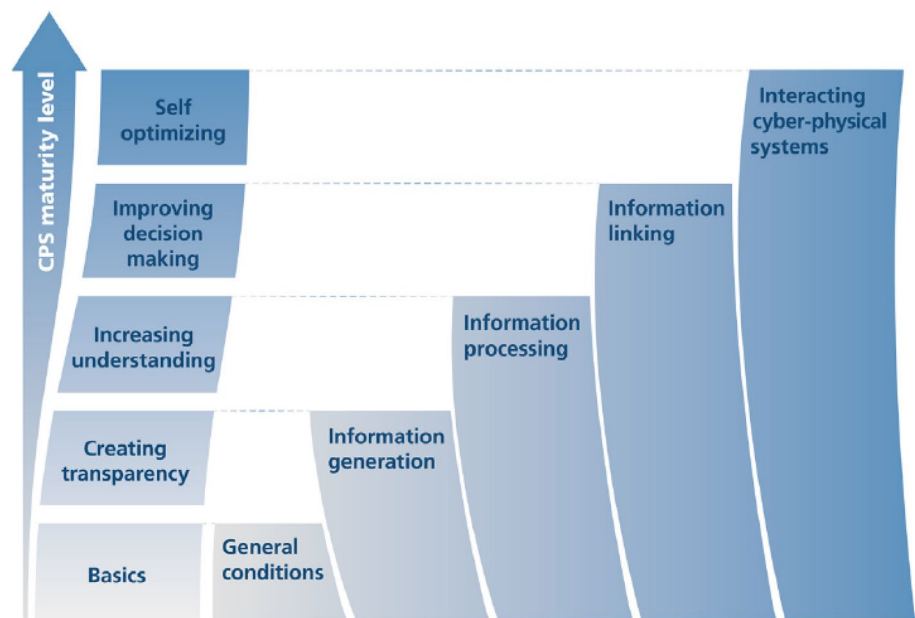
Fig. 2 CPS maturity model [14]

Fig. 3 Energy consumption in the textile industry [10]

processing, and analysis of the cyber world [14, 15]. The physical world, which is the actual field of a CPS, consists of a process machinery that provides integrated functions based on the IIoT convergence technology, devices for data collection, and production equipment. As the information required for production is collected and a platform for an efficient manufacturing environment is created, the entire production process of a physical world can be controlled. Therefore, it is possible to optimize and affect not only the process, but also the subsidiaries [14].

Figure 2 shows the CPS maturity model. The CPS maturity level increases through efforts to improve understanding, accumulate data, and improve decision making. Efficiency improvement and production process optimization are achieved through the application of a CPS [14]. Understanding the physical world is essential for the implementation and application of a CPS, and a cyber world can be effective if it is constructed based on such an understanding [15].

Among CPSs, the CPES is focused on improving the energy efficiency and optimizing processes and production using methodologies, such as mathematical modeling, data analysis techniques, and simulations based on the energy-related information of actual factories obtained through data collection, processing, and analysis [5, 16].

The components of the cyber world are analyzed based on an understanding and an analysis of the site (i.e., the physical world). The physical world can be understood through various data analysis techniques, such as modeling, estimation, and generalization. Data is collected using IIoT devices, and this process can be automated. The CPES constructed through the proposed method improves the CPS maturity level of decision making.

2.2 Energy-Related Status of Dyeing and Finishing Shops

The energy consumption of the textile industry in South Korea in 2015 was 1634.1×10^9 kcal, and that of the local dyeing and finishing shops was 682.5×10^9 kcal (Fig. 3). Thus, these factories represented 42% of the energy consumed by the entire textile industry [10]. When divided by the process, dyeing and finishing polyester consumed 96,723 toe during preprocessing, 204,193 toe during the dyeing process, and 161,205 toe during the finishing process. The dyeing process consumed as much as 44.2% of the total energy required by the dyeing and finishing shops [17].

Energy-related problems in dyeing and finishing shops can be divided into policy problems and technical obstacles. In the case of policy problems, no means can be used to quantitatively predict the potential and effect of improving the energy efficiency. Moreover, efforts to develop standard

energy-saving modules for saving policies without quantitative goals are insufficient. While support policies have been focused on developing and deploying new equipment, the benefits are limited for small- and medium-sized companies because of their incapability to invest [17].

In the case of technical obstacles, the speed of technology development is slow for low-energy equipment that can shorten processes. For small companies, determining the optimal dyeing conditions considering the characteristics of each yarn is difficult, and the results of the work differ in each trial because of ad-hoc decisions by a human operator [17, 18]. These obstacles cause defective products in the dyeing process, and approximately 15–20% of the products require repeated-dyeing. This process increases energy consumption, lowers the product quality, and causes differences in color [17, 19].

This study developed a cyber physical energy system to improve the energy efficiency of small- and medium-sized companies through the collection and utilization of manufacturing big data instead of investing in costly equipment. In addition, methods to reduce the energy consumption by lowering the repeated-dyeing rate were considered. The application and effects of these measures can be predicted through a system approach.

2.3 Machine Learning Techniques

2.3.1 Synthetic Minority Over-Sampling Technique

The synthetic minority over-sampling technique (SMOTE) is an over-sampling methodology to bring the ratio between the minority and majority classes of an imbalanced dataset to an appropriate level. This methodology is applied to successive datasets, generates a random number of [0, 1], and multiplies it by the difference between the randomly selected data and the kth nearest neighbor to replicate a new sample. The populate function, which is the most important function of SMOTE, can be expressed by the following pseudo code: [20].

```
Algorithm Populate(N, i, nnarray)
while N ≠ 0
    Choose a random number nn between 1 and k. This step chooses
    one of the k nearest neighbors of i.
    for attr ← 1 to numattrs
        Compute: dif = Sample[nnarray[nn]][attr] − Sample[i][attr]
        Compute: gap = random number between 0 and 1
        Synthetic[newindex][attr] = Sample[i][attr] + gap ∗ dif
    endfor
    newindex++
    N = N - 1
endwhile
return (∗ End of Populate. ∗)
```

2.3.2 Artificial Neural Networks

Artificial neural networks (ANNs) are defined as "massively parallel interconnected networks of simple elements and their hierarchical organizations which are intended to interact with the objects of the real world in the same way as biological nervous system do" [21]. ANNs have outstanding speed in processing massive parallelism, excellent learning and adaptability, robustness against defects and failures, and wide applicability [22, 23].

ANNs have been used in product design, process planning, scheduling, process modeling and control, monitoring, and diagnosis. As an algorithm with a wide applicability, an ANN has been used at actual manufacturing sites with a demonstrated excellence in performance [23]. In this study, an ANN was used to provide process parameters for the dyeing process as regression models and predict repeated-dyeing in advance as a binary classification model.

3 Cyber Physical Energy System for Dyeing Process

3.1 Problem Description

This section defines the problems of dyeing and finishing shops that are to be addressed with CPES. As noted in Sect. 2.1, understanding the actual physical world is essential, and solution approaches must be set for the implementation and the application of the CPES. In addition, problems that can be solved with CPES application and those that would not be significantly affected must be clearly distinguished.

Dyeing and finishing shops, which are essential elements of the textile industry and require a large amount of energy, have some large-scale workplaces, but are mostly operated in the form of small- and medium-sized enterprises. In this case, fabric is provided by the buyer requesting dyeing and finishing, and the dyeing and finishing shop does not pay any cost related to the fabric purchase. Therefore, the costs incurred by the dyeing and finishing shop are mostly energy related. Previous studies showed that these costs are usually incurred during the dyeing process.

As shown in Fig. 4, the buyer provides the dyeing and finishing shop with the fabric and information on the fabric, required dyeing method, and color. Based on the provided fabric and information, the laboratory of the dyeing and finishing shop cuts a small amount of the fabric and uses it in experiments. The laboratory transfers the process instructions containing various process parameters, such as temperature change, reel velocity, and steam, derived from the experiment to the onsite operator through a process instruction document. The onsite operator adjusts the
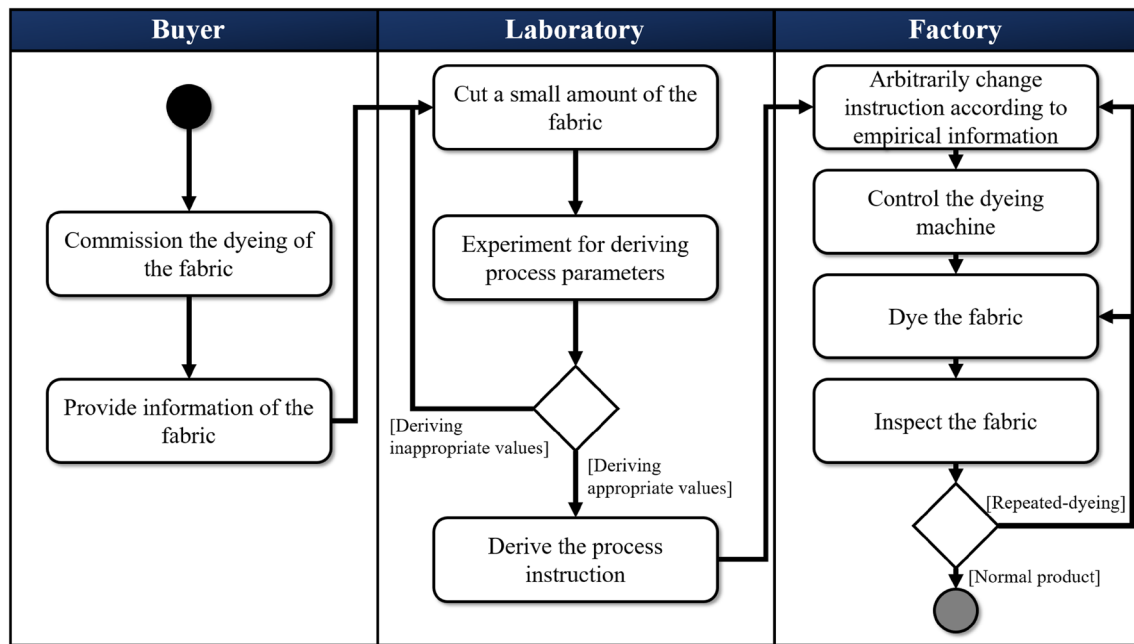
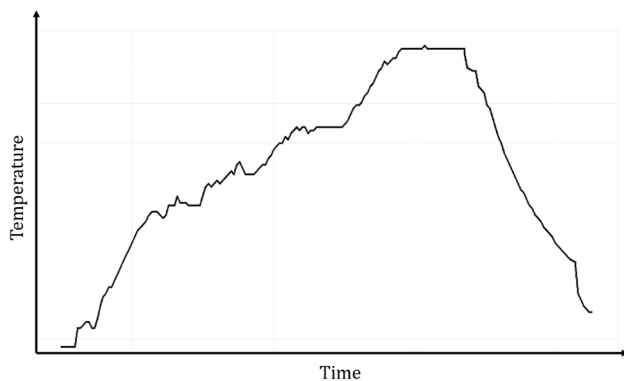**Fig. 4** Activity diagram of the traditional dyeing process scenario

Fig. 5 Temperature curve of the dyeing process

process instructions based on their empirical information and uses these to control the dyeing machine. The operator also continuously changes the control for the gap between the process instructions and the site work. Repeated-dyeing occurs, and the process is performed again if the dyeing results of the dyed fabric are not satisfactory.

For the dyeing process, a high repeated-dyeing rate is an obstacle to increasing energy efficiency, and inefficient processes affect this process. The machines used in the actual laboratory differ from the huge dyeing machines onsite, which produces a significant difference in performance. Therefore, onsite operators must adjust the machines based on their empirical information because incorrect process parameters may result in repeated-dyeing. Upgrading equipment or improving the process efficiency is practically

difficult because of the high costs, shortage of manpower, and other factors.

The temperature data, which are the most important onsite information, are controlled in the order of rise–hold–decline (Fig. 5). This method is mainly used when a non-continuous dyeing machine is used to produce small quantities of various products. The dyeing machine is heated to the maximum temperature and maintained there for a certain time before being cooled again. Field research confirmed that this section is a key factor to understanding the dyeing process.

The process efficiency must be increased and the repeated-dyeing rate must be reduced to increase the energy efficiency of the dyeing process through the construction of a CPES that uses IIoT and manufacturing big data. A difference in the actual onsite conditions is found when process instructions are derived based on an experiment using a small amount of fabric in the laboratory. Therefore, the big data of the actual site must be used to provide process parameters for the dyeing process. Cost reduction and energy efficiency improvement must be achieved not according to the empirical information of the operators, but through understanding the rise–hold–decline of the temperature during the dyeing process based on the manufacturing big data.

3.2 CPES Architecture and Data Model

This section describes the architecture based on which components are used for the CPES for the dyeing process and interactions that are defined between these components.
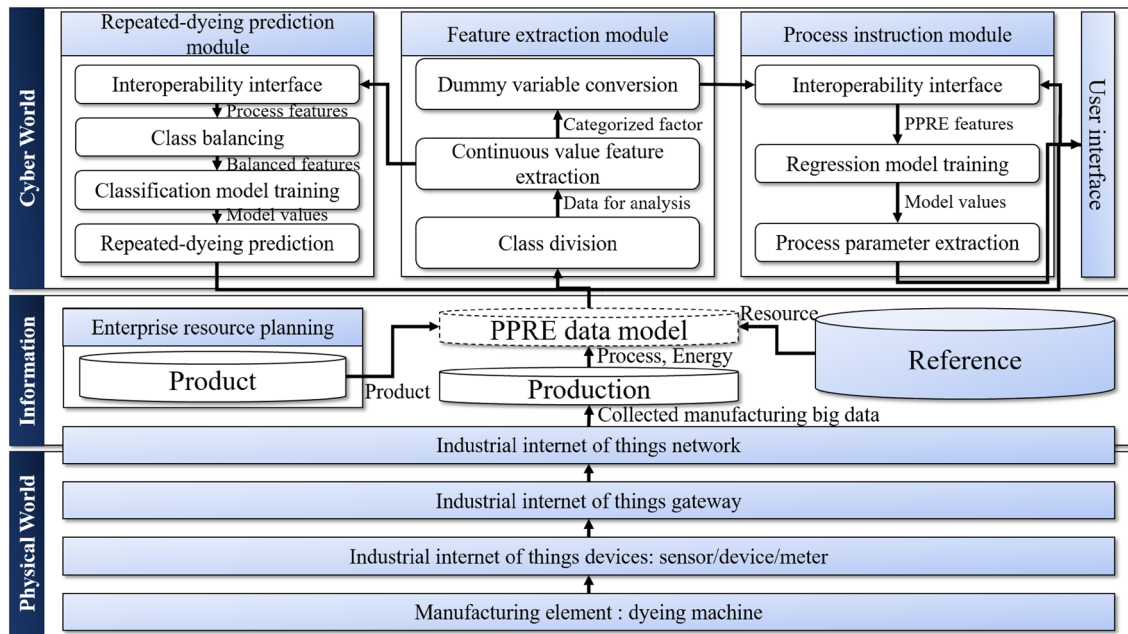
Fig. 6 Cyber physical energy system architecture of the dyeing process

The CPES improves energy efficiency by sharing and using information to connect the physical world and cyber world.

The data used in the CPES of the dyeing process are defined. The time series data collected by the IIoT devices or product data provided by the buyer are defined to apply the techniques related to the manufacturing big data in the proposed CPES. Factors to be considered for preprocessing as variables are also described.

3.2.1 CPES Architecture for the Dyeing Process

The CPES architecture is designed to increase the energy efficiency through sharing and utilization of information gathered in the actual field (Fig. 6). Information is collected through IIoT devices attached to onsite process machines and transmitted through controllers. This information is stored and indexed in databases, such as the system, manufacturing big data, and production databases. The onsite data collected through the IIoT devices are stored in the production database. The Product, Process, Resource, Energy (PPRE) data model is the abstract data repository that retrieves data from three separate databases (i.e., production, reference, and product databases) in enterprise resource planning (ERP) and indexes them for use within the CPES. Applications in the cyber world operate based on the collected data based on this PPRE data model.

The cyber world consists of a feature extraction module that extracts variables based on the PPRE data model, buyer order information, and product property information; a process instruction module that uses these variables and

provides process parameters; and a repeated-dyeing prediction module that determines the repeated-dyeing possibility for the onsite dyeing process. The module, sensor, and information repository in the CPES perform processes, such as collecting, indexing, sharing, processing, and utilizing information, and provide onsite operators with information on improving the energy efficiency.

Unlike the existing approach of process parameters through experiments, the variables are extracted using the manufacturing big data collected from the site. The learning model is then used to instruct the process parameters. Once the buyer inputs the required order and fabric information into the learning model through the user interface, the process instruction module derives the process parameters for the dyeing process, and the process parameters are transmitted to the onsite operators controlling the dyeing machines. In this case, the onsite operators need not continuously adjust the process parameters based on their empirical information because the process parameters are derived from the information of the actual site instead of the information obtained from the experiments in the laboratory. Figure 7 illustrates the activity diagram showing the advanced scenario of the dyeing process when using the CPES to improve the energy efficiency.

During the process, the repeated-dyeing prediction module derives the repeated-dyeing possibility for each section of the rise–hold–decline of the temperature in the dyeing process with the binary classification algorithm and sends the information to the onsite operators. Even if the repeated-dyeing possibility increases because of process parameter
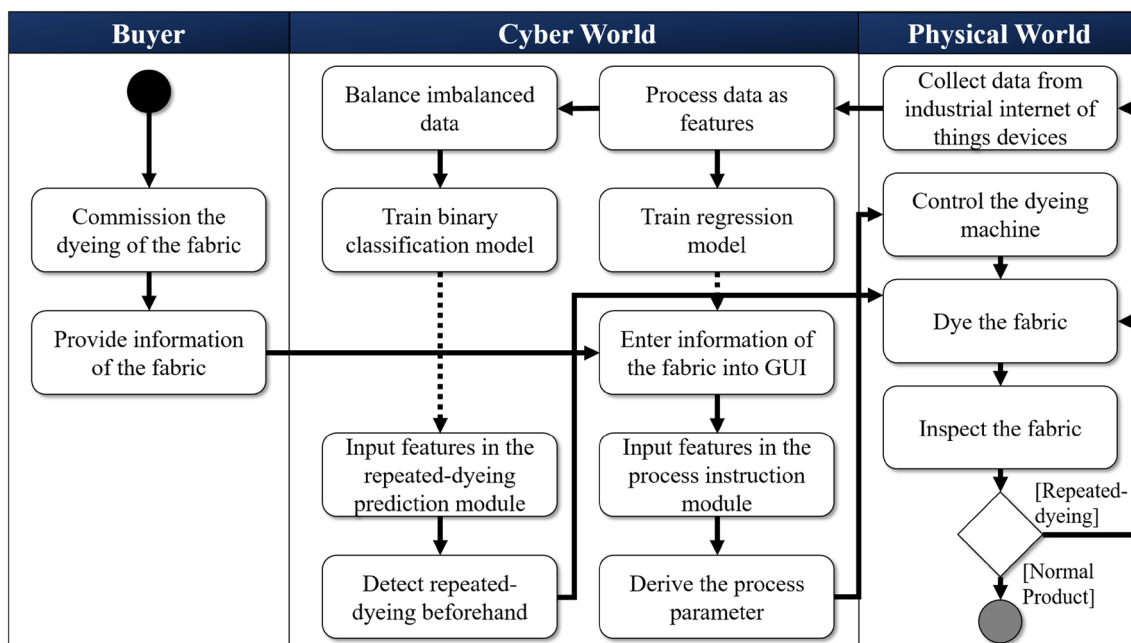
Fig. 7 Activity diagram of the advanced scenario of the dyeing process with the cyber physical energy system
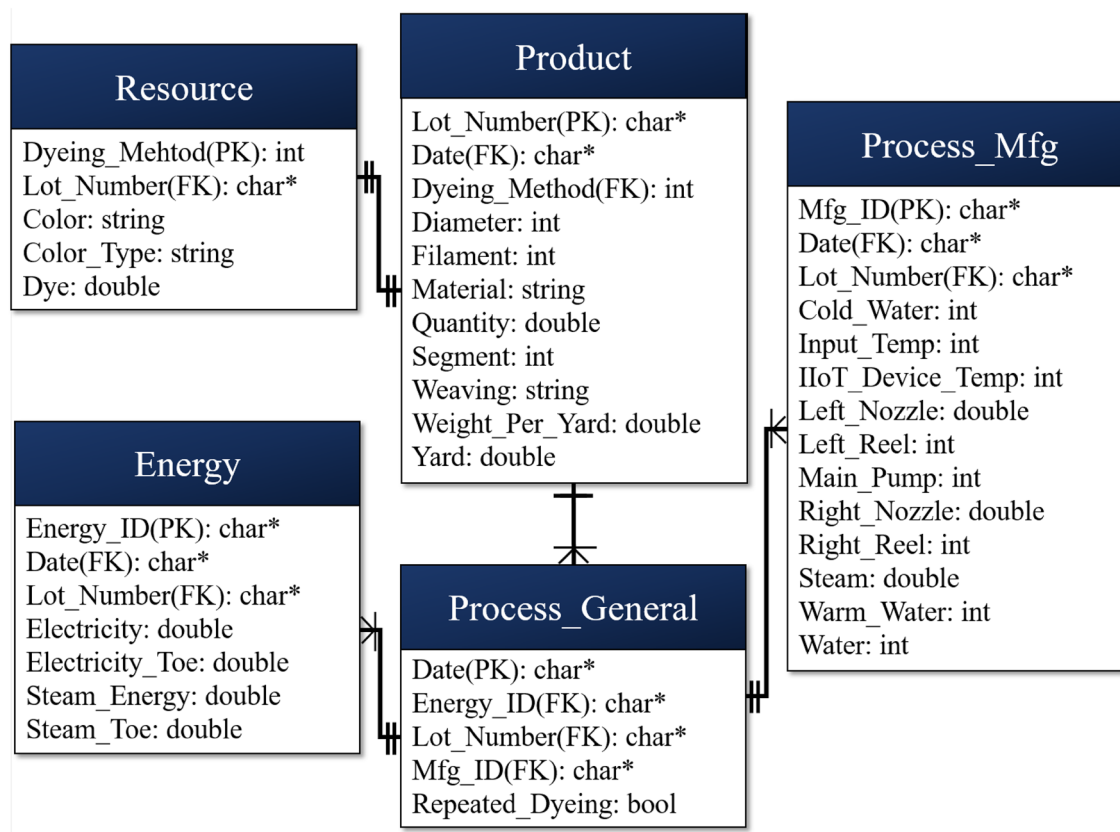

Fig. 8 View of PPRE data model

errors or process problems, the operators can handle the dyeing process based on the repeated-dyeing possibility according to the manufacturing big data instead of their uncertain own experience.

3.2.2 PPRE Data Model for the CPES

An abstract data model is needed to retrieve and use data from three different databases in the architecture described in Sect. 3.2.1. The model is summarized in this section based on the keyword PPRE. The entire model view is shown, and the schemas of each component are described. Figure 8 shows the PPRE data model for the CPES of the dyeing process used herein.

The PPRE data model consists of five entities, each referring to data retrieved from the three databases through query statements. The product entity retrieves the texture and order data from the ERP database. The resource entity is obtained from the reference database with specific data of the dyeing method. The other three entities are taken from the production database for the manufacturing big data collected through the IIoT environment.

Table 1 presents the schema of the product entity. These elements are data that can be collected when the dyeing and finishing shop receives a dyeing order from a buyer and are retrieved from the ERP database. *Lot_Number* is used as the ID of the schema and to divide the data samples. Data, such as *Filament*, *Segment*, and *Weaving*, which significantly affect the dyeing quality, are also collected. These data are so varied that they make it difficult to provide process parameters through experiments.

Table 2 presents the resource schema in the PPRE data model. The elements in Table 2 can be corresponded when a dyeing order is received from a buyer. These elements are retrieved from the reference database of the company itself. The data are mainly corresponded when the order is received; hence, they usually cannot be modified according to the needs of the site. These data include information that do not affect the actual work results along with the order data requested by the buyer, including the *Color* and the *Dyeing_Method*, and the material property data of the fabric to be dyed, including the *Segment* and the *Quantity*.

Table 3 presents the schema in the process_general schema, while Table 4 presents the schema in the process_mfg schema. These data are related to the onsite processes

Table 1 Schema of product entity in PPRE data model [9]

Element	Unit	Type	Definition
Lot_Number	Code	Characters	Lot number of fabric to be dyed
Diameter	Inches	Integer	Diameter of yarn of fabric to be dyed
Material	Code	String	Name of fabric to be dyed
Filament	Inches	Integer	Density level of fabric to be dyed
Quantity	Kilograms	Double	Weight of fabric to be dyed
Segment	Number	Integer	Number of segments of fabric to be dyed
Weaving	Number	String	Pattern type of fabric to be dyed
Weight_Per_Yard	Grams per yard	Double	Weight per yard of fabric to be dyed
Yard	Yard	Double	Yard of fabric to be dyed

Table 2 Schema of resource entity in PPRE data model [9]

Element	Unit	Type	Definition
Color	Code	String	Color of fabric after dyeing that has been requested by the buyer
Color_Type	Code	String	Type of color of fabric after dyeing that has been requested by the buyer (dark, light, white, etc.)
Dye	Grams	Double	Amount of dye to be input to the dyeing process
Dyeing_Method	Code	Integer	Dyeing method to be used in the dyeing process that has been requested by the buyer

Table 3 Schema of process_general entity in PPRE data model [9]

Element	Unit	Type	Definition
Date	Date	String	Date on which the process is conducted
Repeated_Dyeing	True or False	Boolean	Occurrence of repeated-dyeing during the dyeing process

Table 4 Schema of process_mfg entity in PPRE data model [9]

Element	Unit	Type	Definition
Cold_Water	L	Integer	Cold water used in the dyeing process
Input_Temp	°C	Integer	Input temperature to operate the dyeing machine
IIoT_Device_Temp	°C	Integer	Temperature collected from the IIoT devices
Left_Nozzle	kgf/cm^3	Double	Injection pressure of the left nozzle that injects the dye
Left_Reel	rpm	Integer	Velocity of the left reel
Main_Pump	rpm	Integer	Velocity of the main pump used in the dyeing process
Nozzle	kgf/cm^3	Double	Injection pressure of the left and right nozzles that inject dye
Reel	rpm	Integer	Velocity of the left and right reels
Right_Nozzle	kgf/cm^3	Double	Injection pressure of the right nozzle that injects the dye
Right_Reel	rpm	Integer	Velocity of the right reel
Steam	kg/h	Double	Steam used in the dyeing process
Warm_Water	L	Integer	Warm water used in the dyeing process
Water	L	Integer	Room-temperature water used in the dyeing process

Table 5 Schema of energy entity in PPRE data model [9]

Element	Unit	Type	Definition
Electricity	kw/h	Double	Electrical power used in the dyeing process
Electricity_Toe	toe	Double	Amount of electrical energy used in the dyeing process
Energy	toe	Double	Total amount of energy used in the dyeing process
Steam_Energy	kg/h	Double	Amount of steam energy used in the dyeing process
Steam_Toe	toe	Double	Steam used in the dyeing process

collected through IIoT devices, such as watt-hour meters, water meters, and temperature sensors, during the actual dyeing process. These data represent the actual onsite operation based on the process parameters derived from the experiments in a laboratory using data, such as the material properties in Table 1 and order requests. As noted in Sect. 3.1, onsite operators use their empirical information to adjust the operation instead of following the exact process parameters provided by the laboratory.

Table 5 presents a schema of the energy entity containing information related to the energy used in the dyeing process. Steam and electric energy are the main energy sources required by the dyeing process. Such information is collected from steam meters and watt meters.

3.3 Cyber World Configuration of the CPES

This section describes the modules that constitute the cyber world of the CPES. Each module replaces the existing laboratory experiments and experience of onsite operators by applying machine learning techniques to the IIoT-based manufacturing big data and reference data.

3.3.1 Feature Extraction Module

After various data are collected and stored in the CPES scenario, the feature extraction module preprocesses the data for use in the analysis modules (Fig. 9). Various data, such as onsite data, order information, and product information, cannot be directly used; hence, they must be converted into variables based on the understanding of the dyeing process for utilization.

The rise–hold–decline sections need to be considered separately in the dyeing process that uses non-continuous dyeing machines; thus, the feature extraction module divides those three sections and extracts variables for consideration.

The extracted variables are stored and used as a training set for other modules or when variables are preprocessed to instruct process parameters and determine the repeated-dyeing possibility with the constructed learning model. As the first step of the procedure of the feature extraction module, the module checks if the data are interrupted because of communication problems.

The Product and Resource data are replaced by nominal variables if the data are normally collected. At this step, continuous variables, such as diameter in the Product data and dye in the Resource data, are used as variables without additional preprocessing. These variables are used only in the regression model of the process instruction module; therefore, they are replaced by dummy variables.

In contrast, in the case of the Process and Energy data, the module divides the section as shown in Fig. 10 and extracts the feature based on it. Table 6 presents the extraction of the feature using the value for each data. The nozzle pressure

Table 6 Feature extraction method by data

Source	Data	Type	Method
Product	*Diameter*	Continuous	Extract data itself as a variable without preprocessing
	Material	Continuous	Extract data itself as a variable without preprocessing
	Filament	Category	Extract dummy variable after nominal variable conversion
	Quantity	Continuous	Extract data itself as a variable without preprocessing
	Segment	Continuous	Extract data itself as a variable without preprocessing
	Weaving	Continuous	Extract data itself as a variable without preprocessing
	Weight_Per_Yard	Continuous	Extract data itself as a variable without preprocessing
	Yard	Continuous	Extract data itself as a variable without preprocessing
Resource	*Color*	Category	Extract dummy variable after nominal variable conversion
	Color_Type	Category	Extract dummy variable after nominal variable conversion
	Dye	Continuous	Extract data itself as a variable without preprocessing
	Dyeing_Method	Category	Extract dummy variable after nominal variable conversion
Process_general	*Repeated_Dyeing*	Category	Notate binary class as 0 and 1
Process_mfg	*Cold_Water*	Continuous	Extract cumulative value as a variable
	Left_Nozzle	Continuous	Extract cumulative value as a variable
	Left_Reel	Continuous	Extract average value as a variable
	Main_Pump	Continuous	Extract average value as a variable
	Right_Nozzle	Continuous	Extract average value as a variable
	Right_Reel	Continuous	Extract cumulative value as a variable
	Steam	Continuous	Extract cumulative value as a variable
	Warm_Water	Continuous	Extract cumulative value as a variable
	Water	Continuous	Extract cumulative value as a variable
Energy	*Electricity*	Continuous	Extract cumulative value as a variable
	Electricity_Toe	Continuous	Extract cumulative value as a variable
	Energy	Continuous	Extract cumulative value as a variable
	Steam_Energy	Continuous	Extract cumulative value as a variable
	Steam_Toe	Continuous	Extract cumulative value as a variable

and the reel velocity are excluded because the sum of the left and right data forms an integrated value.

The temperature data of dyeing machines have complex patterns that are difficult to formulate; hence, additional preprocessing for variable extraction is required. The least-squares method is used as a preprocessing rule. The rising and declining accelerations are estimated as variables based on the slope of the regression line, which is represented by the dotted line in Fig. 10 [12]. Table 7 presents the dyeing curve preprocessing method.

The complex temperature rising and declining patterns are generalized using one variable estimated through preprocessing. In addition, the variables in the process_mfg entity are extracted according to this section, and the product and resource data are converted into nominal data to become dummy variables. The manufacturing big data are generalized as parameters for each section using the cumulative, maximum, and average values.

The dataset is preprocessed if the process parameters for the dyeing process and the repeated-dyeing possibility must be determined. The data that did not cause repeated-dyeing

are extracted and preprocessed as a training set upon the request of the process instruction module.

3.3.2 Process Instruction Module

The process instruction module instructs the process parameters based on a multiple regression model. Figure 11 shows the procedure employed by the process instruction module.

The first step of the procedure is a request for training samples from the feature extraction module. The module performs normalization after receiving the training sample, in which all variables are normalized to adjust biased data for instructing process parameters using the regression models. Thus, the corresponding variable is prevented from controlling the result.

The steps for training the multiple regression model in the procedure are designed to extract a meaningful output through a repetitive experiment. A regression model is constructed for each output variable based on the normalized preprocessed variables. Such multiple regression models derive process parameter variables one by one according
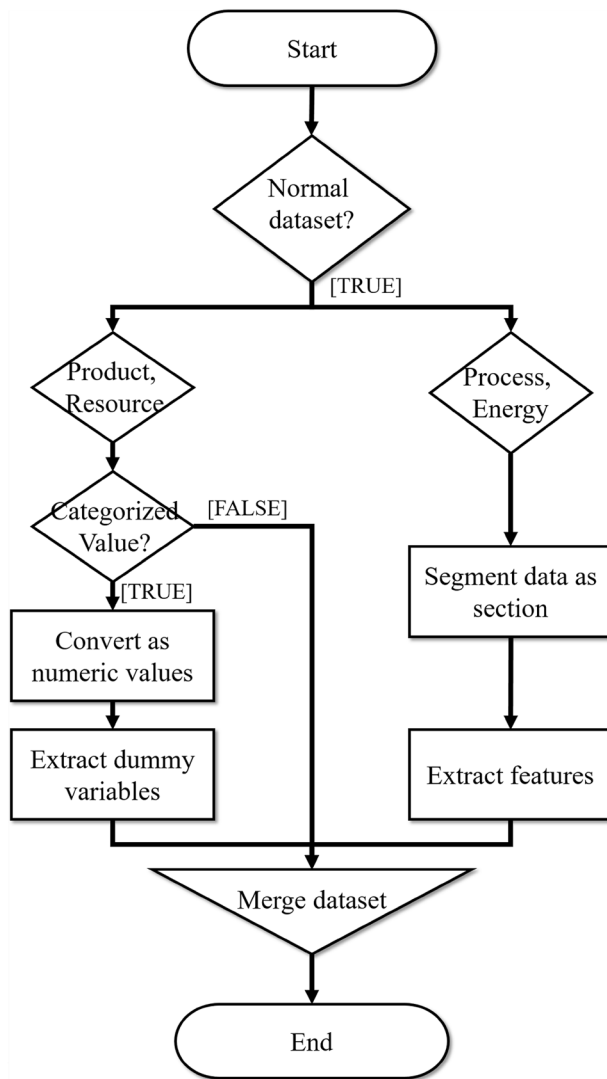
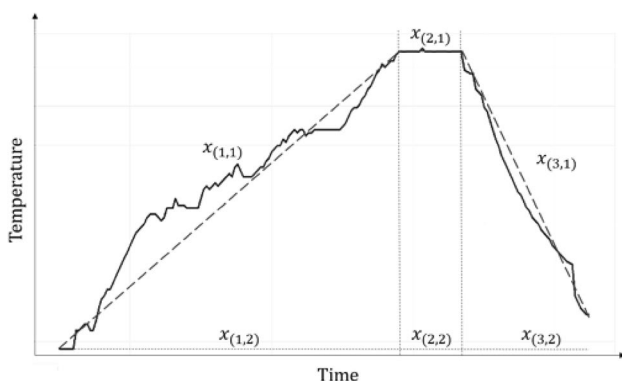**Fig. 9** Procedure employed by the feature extraction module

Fig. 10 Variable extraction on the temperature curve [12]

Table 7 Preprocessing method for dyeing curve

Data	Notation	Method
Increasing_Slope	$\widehat{x_{(1,1)}}$	$\dfrac{\sum_{section1}\left(tmp_i - \overline{tmp}\right)\left(time_i - \overline{time}\right)}{\sum_{section1}\left(time_i - \overline{time}\right)}$
Increasing_Time	$x_{(1,2)}$	*Time of Section 1*
Maximum_Temperature	$x_{(2,1)}$	*Max (Section 2)*
Max_Temper_Time	$x_{(2,2)}$	*Time of Section 2*
Decreasing_Slope	$\widehat{x_{(3,1)}}$	$\dfrac{\sum_{Section\ 3}\left(tmp_i - \overline{tmp}\right)\left(time_i - \overline{time}\right)}{\sum_{Section\ 3}\left(time_i - \overline{time}\right)^2}$
Decreasing_Time	$x_{(3,2)}$	*Time of Section 3*

correlation analysis and for adjusting epoch to avoid overfitting compared to the multivariate model. The multivariate model also requires a considerable learning time from the large number of change types and has a low prediction accuracy.

The input variables of these models are derived from the data and can be collected before work, product, and resource extracted as dummy variables. The output variables are derived from the data and can be collected during the process, process, and energy data extracted in every interval.

The set of outputs derived by inputting new order information into the learned regression model is inverted using the normalization in the opposite manner. A combination of process parameter values is then derived. The process instructions are set based on this derived combination.

3.3.3 Repeated-Dyeing Prediction Module

The repeated-dyeing prediction module collects the data of the dyeing process currently in operation and predicts the repeated-dyeing possibility for each section of the rise–hold–decline process of the temperature. Figure 12 shows the procedure of the repeated-dyeing prediction module. For prior detection, this module uses binary classification models constructed for each section.

After the existing data are preprocessed using the feature extraction module and the preprocessed data are imported as the first step of the procedure, the module sorts variables by each section for training the classification models. These variables are features extracted from the process data, which are continuous values.

The next step is to oversample the negative class of the feature to the positive class and normalize the elements in both classes. These two steps are preprocessing steps to reduce the effects of certain values and improve the accuracy of the classification models.

Next, binary classification models are constructed using the adjusted samples. The input variables of these
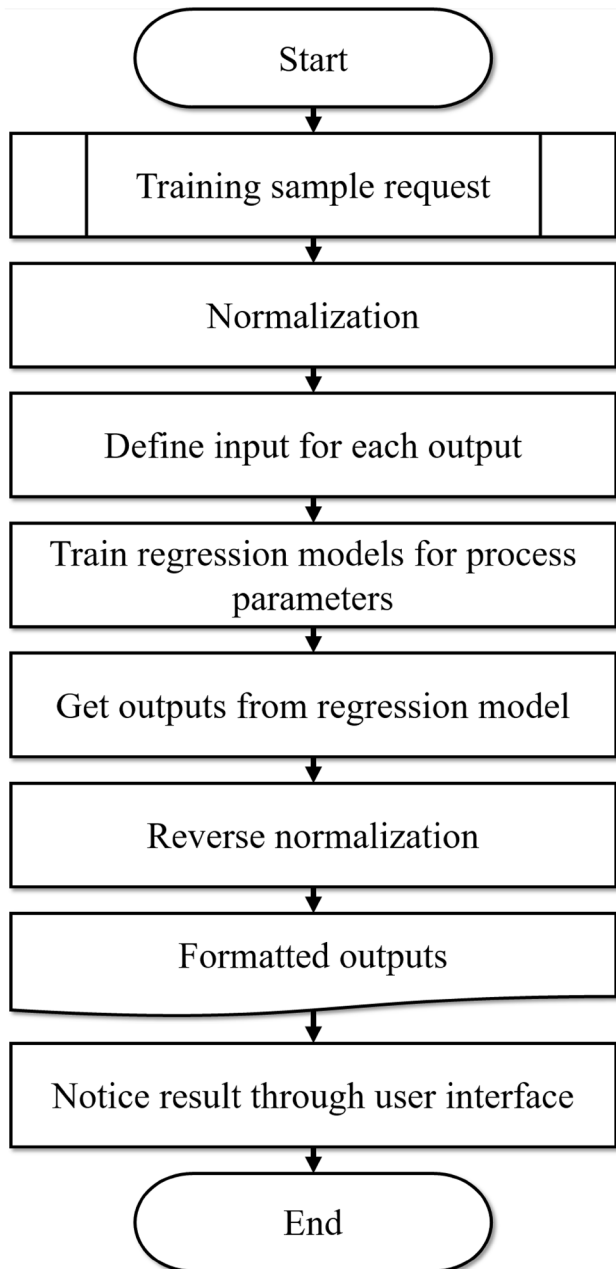
to the input variables. Multiple models are trained because they are more efficient for selecting each input based on a

Fig. 11 Procedure employed by the process instruction module

Flowchart (Fig. 11): Start → Training sample request → Normalization → Define input for each output → Train regression models for process parameters → Get outputs from regression model → Reverse normalization → Formatted outputs → Notice result through user interface → End

Fig. 12 Procedure employed by the repeated-dyeing prediction module

Flowchart (Fig. 12): Start → Request data processing → Normalization → Sort variables by each section → Balance binary class → Train classification models for prediction → Get input for each section → Request real-time data processing → Predict repeated-dyeing → End

classification models are process variables, while the output variables are the binary class of the *Repeated_Dyeing* variable.

During the actual dyeing process, a classification model is used for each section. The operators should be notified if a negative class with a high repeated-dyeing possibility is determined. The operators are aware of the high repeated-dyeing possibility before work is completed; hence, they can prepare for it. The operators at the actual site must continuously check whether repeated-dyeing will occur or not. Such waste can be prevented by improving the process efficiency.
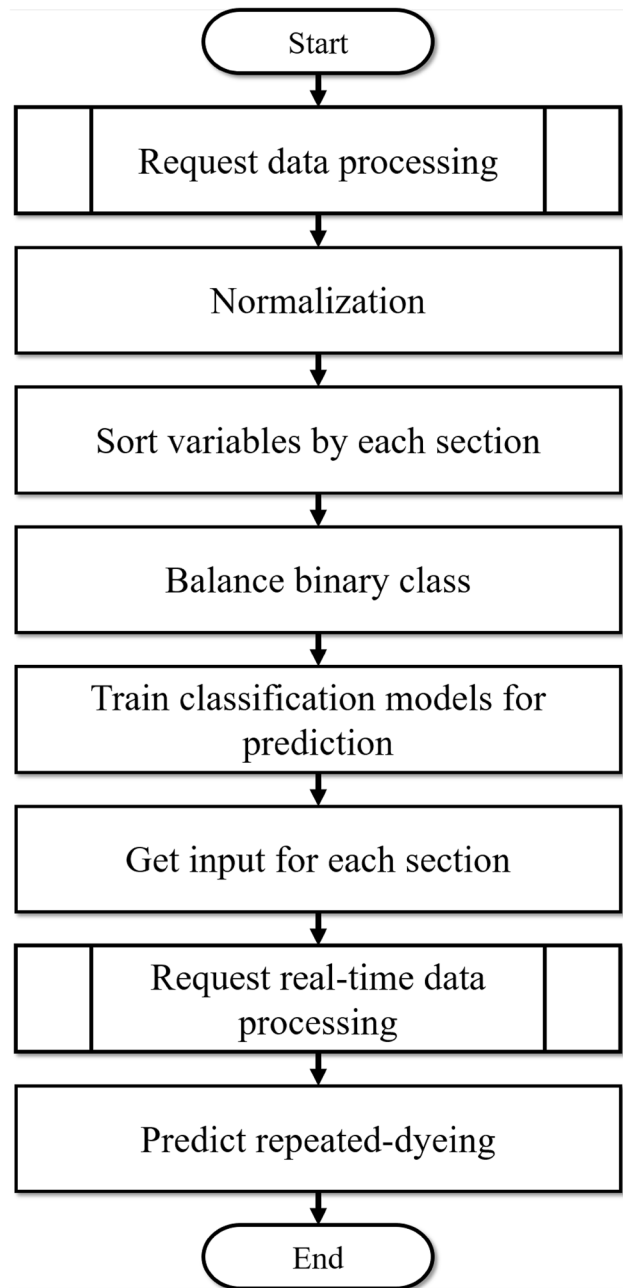
4 Case Study

4.1 Constitution of Environment

The development environment of the modules constituting the cyber world of the CPES and the configuration environment of the IIoT gateway are as described earlier. A rapid machine (SIDC-8200) was selected as the target dyeing machine (Fig. 13). This dyeing machine is non-continuous and goes through the rise–hold–decline process for

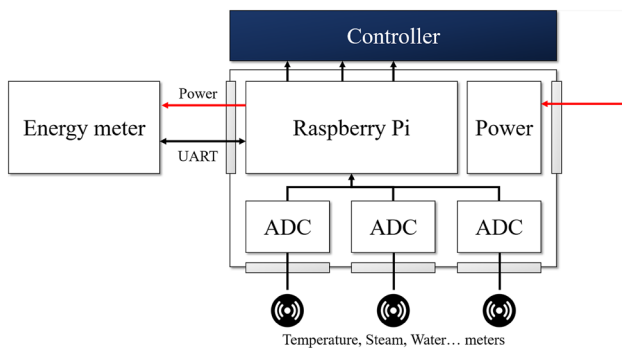**Fig. 13** Dyeing machines in the dyeing and finishing shop

Fig. 14 Industrial internet of things gateway diagram for the case study of the cyber physical energy system

temperature during the dyeing process. Table 8 presents the test environment for this case study.

Scenarios are constructed to verify the validity of the three major parts of the architecture proposed herein. The three parts of the scenario are the main elements that would be considered in an actual application of the CPES and must also be considered before expansion.

In the physical world (i.e., the actual dyeing and finishing shop), IIoT devices are attached to the dyeing machine, and data are collected through the IIoT gateway. Figure 14 shows the process by which IIoT devices are installed in the dyeing machine and data are collected through the IIoT gateway for this case study.

The data from the IIoT devices installed in the physical world are transmitted through transmitters and receivers, converted into digital information, and recorded. Figure 15 shows a screen for recording the temperature curve and the energy consumption in the controller applied to the actual factory.

Table 9 lists the machine learning techniques chosen to implement and verify the procedures and functions of the modules mentioned in Sect. 3.3. An ANN is a machine learning technique with adequate performance when modeling using various inputs. Collecting many samples is difficult because the dyeing and finishing shops process only a small number of products per unit time because of their industrial characteristics and factory size. Therefore, oversampling must be performed to balance the binary class through SMOTE instead of utilizing the under-sampling method.

4.2 Implementation of the CPES Architecture

An abstract database was implemented by querying the collected data from three different databases according to the proposed PPRE data model. The benchmark samples for the process parameters of the dyeing process and the prior
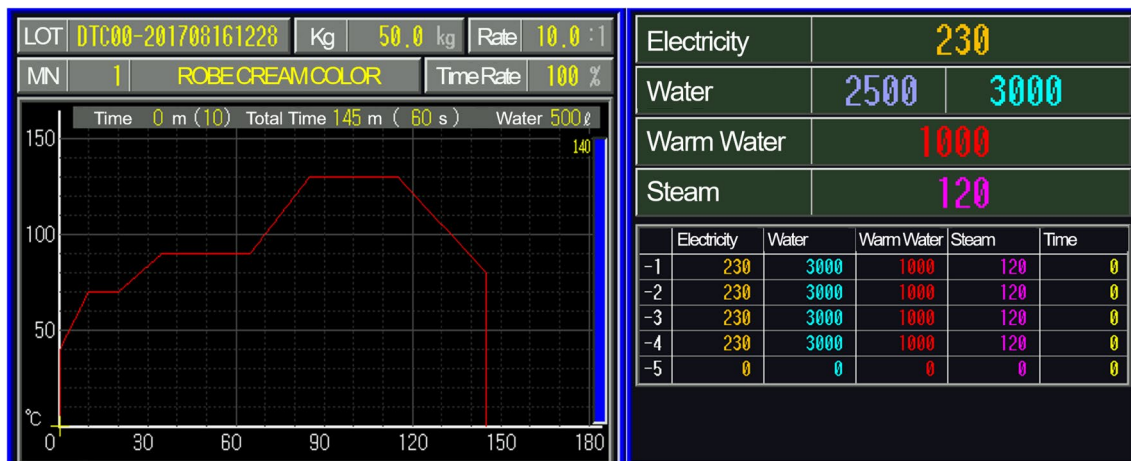

Fig. 15 Human machine interface of the controller of the dyeing machine

Table 8 Test environment for the case study

Item	Contents
CPES module	Window 10 OS 64 bit
	Visual Studio 2015 C#.Net Framework 3.5
	Intel® Core™ i7-6500U CPU @ 2.50 GHz
Dyeing machine	Samil SIDC-8200
IIoT gateway	Broadcom BCM2837 64bit
	ARM Coretex-A53
	1 GB LPDDR2
	Linux/Windows Embedded
IIoT network	RESTful API

detection of the repeated-dyeing possibility were prepared using sample files from the databases, which were preprocessed and balanced with the feature extraction module. Training and testing were performed using the benchmark samples, and the modules were validated.

Figure 16 shows a client that can access the abstract database containing the data according to PPRE data model of the CPES architecture. The manufacturing big data can be accessed if the data collection site and the period are selected. The product and resource data as well as process and energy data of the corresponding case were included. Such data were preprocessed and analyzed through various modules of the cyber world.

Figure 17 shows the class diagram for the actual implementation. The three modules that constituted the cyber world of the CPES architecture were implemented by modular design, and each of these modules can be run independently. In addition, an ANN library was added to construct both regression and classification models.

Figure 18 shows the scatter plots of the features of the benchmark samples for this case study. The dots represent the variables derived by the feature extraction module. The blue diamond dots represent the regular products. The red rectangular dots represent the cases requiring repeated-dyeing. The good product cases were confirmed to form significant areas. Some cases with repeated-dyeing exhibited significant differences that can be considered as outliers.

The general information of the benchmark samples were data extracted from a total of 384 lots and had 82 repeated-dyeing data. This result indicated that the repeated-dyeing rate in these samples was approximately 21.35%. The energy of the dyeing process can be confirmed to be saved if repeated-dyeing did not occur or was detected beforehand.

The process instruction module of the CPES in the case study utilized the ANN technique for the regression model and instructed the process parameters based on such manufacturing big data. The product and resource data of the data model were used as the input variables of the multiple regression models using the ANN methodology in the process instruction module based on the data preprocessed
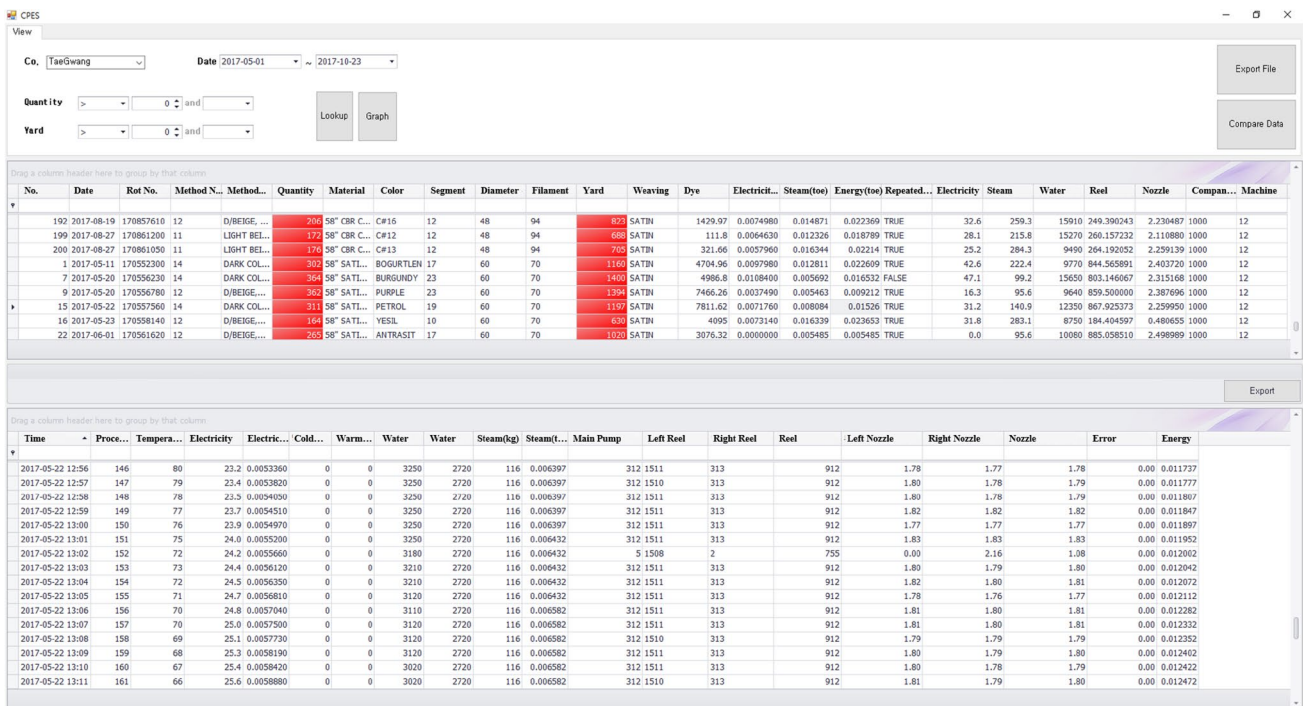

Fig. 16 Implementation of the web-client according to the PPRE data model
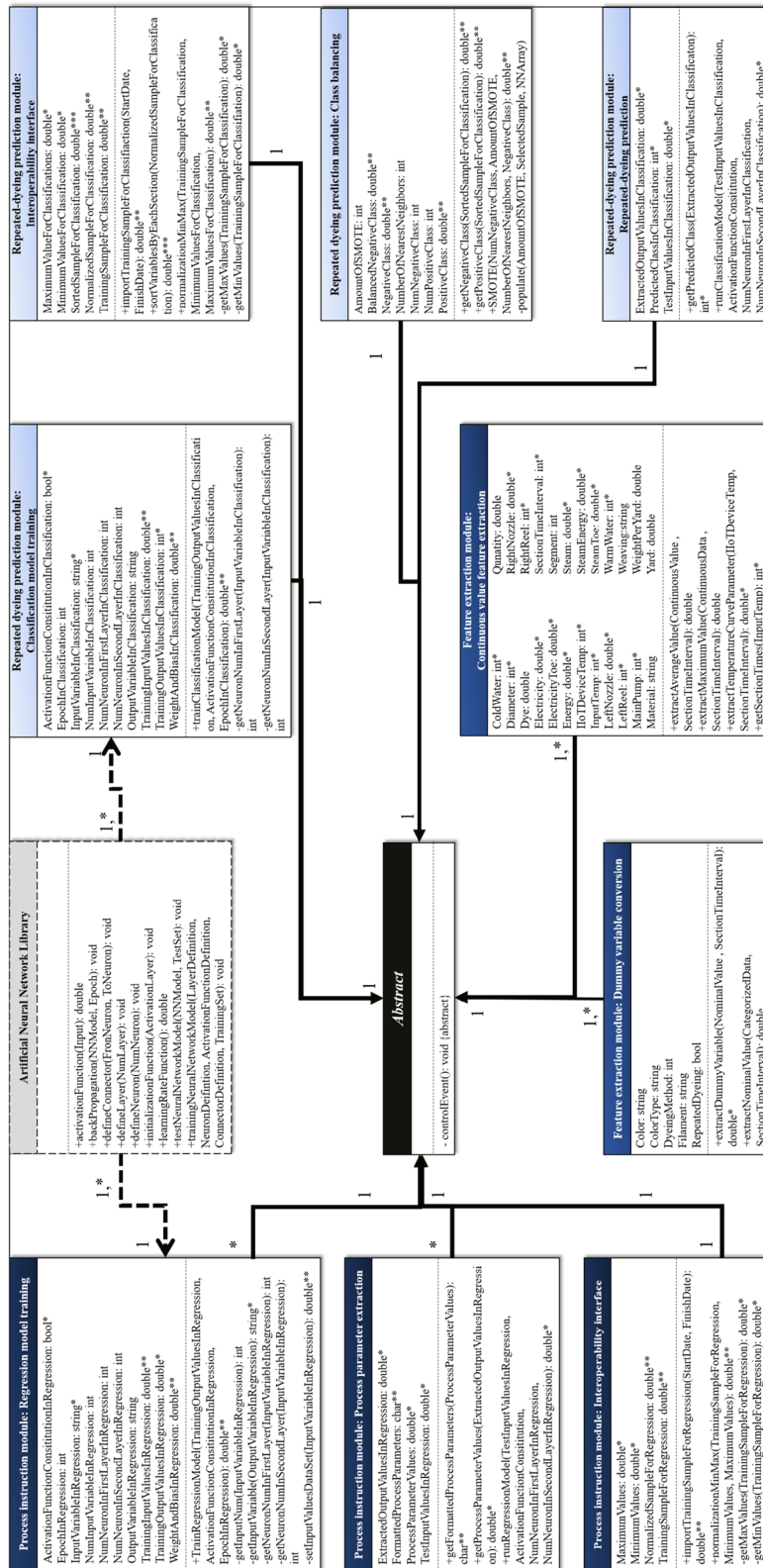
Fig. 17 Class diagram for implementation of cyber world in the cyber physical energy system
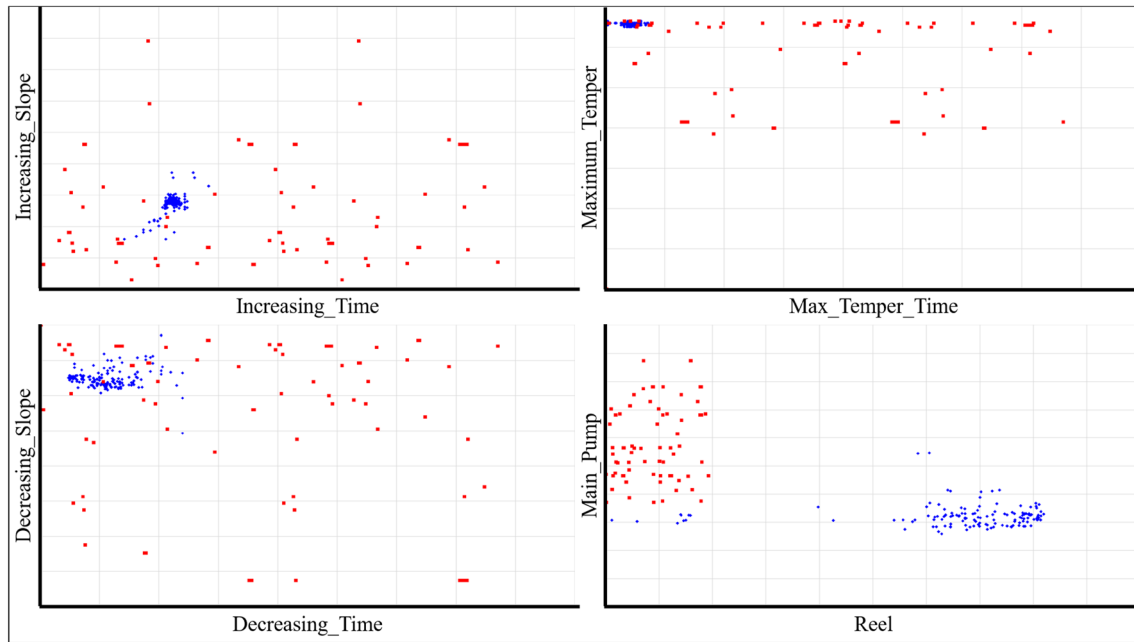
Fig. 18 Scatterplot of features for the case study

Table 9 Machine learning techniques selected in the case study

Item	Contents
Regression modeling technique	Artificial neural network
Binary classification modeling technique	Artificial neural network
Binary class balancing technique	Synthetic minority over-sampling technique

Table 10 Experimental information of regression model used to validate the process instruction module

Item	Contents
Number of training sets	211
Number of test sets	91

through the feature extraction module. The information related to the process were output.

Table 10 presents the information used to validate the process instruction module. A total of 384 samples were used, and 82 samples were found to undergo repeated-dyeing. Only 302 good product samples were selected to match the solution area of the regression model to that of the good products. Learning was also provided. In this instance, the numbers of samples for learning and validation were fixed at 7:3.

The models were set to two hidden layers, and the inputs were defined by selecting variables with a high correlation for each variable. These variables were from the product and

Table 11 Results of the prediction accuracy experiment for the process instruction module

Item	Error rate (%)
Average	0.8687
Maximum	14.3242
Minimum	0.0352

resource data retrieved from the reference and ERP databases. As noted in Sect. 3.3.2, 30 outputs were obtained; hence, 30 regression models with a single output were constructed instead of a multivariate regression model. The results of the prediction accuracy experiment in Table 11 demonstrated that the fitting of the regression model can be performed well.

The possibility of energy reduction was validated based on the product and resource features of the repeated-dyeing samples. Table 12 presents the results. The sum of the energy consumption (toe) with the process parameter that previously caused repeated-dyeing and when the fabric was repeated-dyed was compared with the predicted energy consumption with the process parameter according to the process instruction module.

The ratio of the energy consumption expected with the process instruction module to the total amount of energy involved in repeated-dyeing was calculated for 82 repeated-dyeing samples (Table 12). The average energy consumption was 0.0207391 toe. Therefore, the process parameter from the process instruction module consumed energy, which was only 89.31% of the traditional process. The energy efficiency was improved by approximately 10.69%.

Table 12 Energy consumption comparison between process instruction module and cases with repeated-dyeing

Item	Average energy consumption (toe)
Traditional dyeing process	0.0232215
Process instruction module	0.0207391

Table 13 Experimental information of binary classification model used to validate the repeated-dyeing prediction module

Item	Contents
Number of training sets	
Positive	211
Negative	91
Number of test sets	
Positive	57
Negative	25

Table 13 presents the experimental information for the validation of the repeated-dyeing prediction module. The ratio of the training set to the test set was 7:3. Stratified sampling was performed according to the class of the *Repeated_Dyeing* variable; hence, the repeated-dyeing rates of the training set and the test set were identical. *Repeated_Dyeing* was set as the dependent variable, with 13 variables utilized for each section as inputs. The inputs were variables from the process and energy data. The sum of the other values, such as *Reel, Nozzle, and Energy*, was excluded because the dependencies were too large to utilize in the modeling.

By balancing the binary class step using SMOTE, the number of negative class samples was adjusted to be similar to the number of positive class samples. The binary classification modes were also constructed for each section based on the over-sampled variable samples. Figure 19 shows the results of the experiments for validating the repeated-dyeing prediction module. The receiver-operating characteristic (ROC) curves of each section were drawn in the figure. The area under the curve (AUC) value was confirmed based on these curves (Table 12). The binary classification model in Sect. 1, which was considered the most important section in the actual field, had the largest difference between both classes (Table 13). Through this experiment, it was confirmed that the field operator should be instructed in advance of the repeated-dyeing, which requires 15–20% of the total product from the dyeing process (Table 14). It also confirmed that the performance of the binary classification models using the ANN was of a very high quality, and that the scenario defined can be implemented.
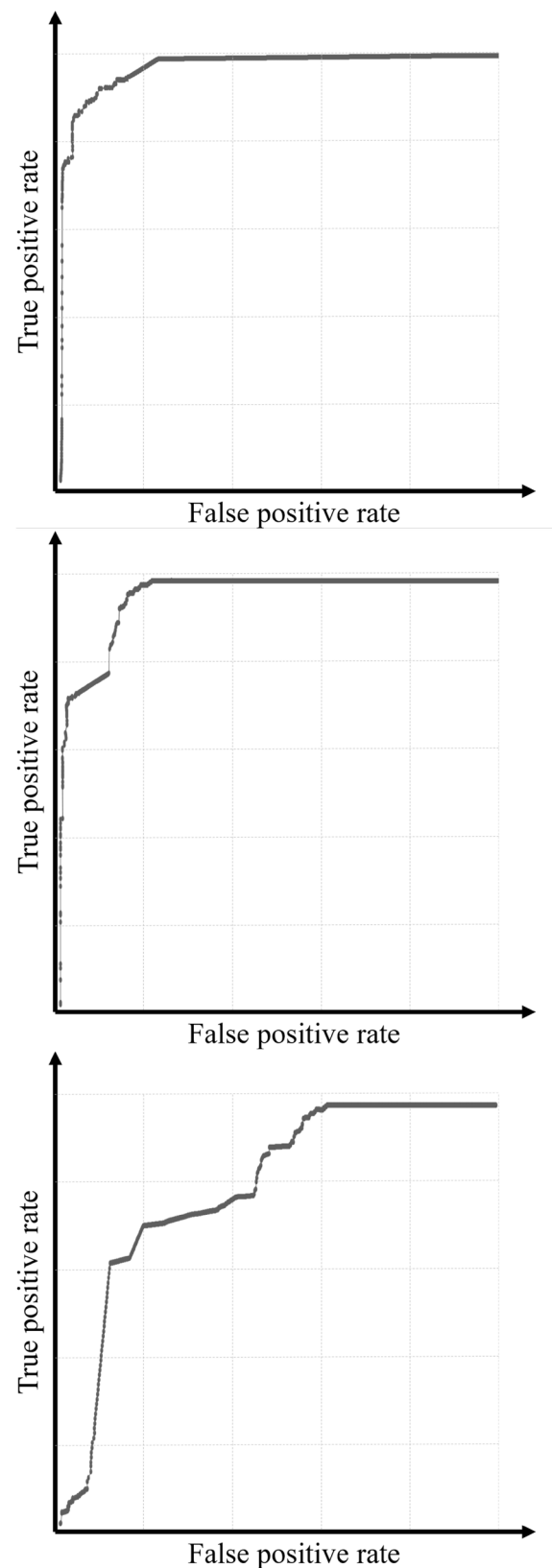

Fig. 19 ROC curves of binary classification models of repeated-dyeing prediction module

Table 14 Experimental results of the repeated-dyeing prediction module

Item	AUC
Section 1	0.982456
Section 2	0.963955
Section 3	0.822329

5 Conclusions

This paper considered the design, structure, and flow of a CPES to improve the energy efficiency of the dyeing process. A PPRE data model for retrieving data from several databases was proposed to store data related to the energy efficiency of the dyeing process. In addition, definitions were provided for the manufacturing big data to be collected for the CPES construction based on understanding of the site. Process parameters for the dyeing process based on the manufacturing big data and predicting repeated-dyeing in advance were derived to improve the energy efficiency.

The preprocessing, utilization, and application of the manufacturing big data collected through the IIoT devices, gateway, and network were examined. Moreover, the procedures employed by the feature extraction module, process instruction module, and repeated-dyeing prediction module were determined. A case study was conducted in which data were collected by installing IIoT devices, a gateway, and a network in an actual dyeing and finishing shop. The validity of the modules was evaluated based on the collected data. A comparison with the benchmark data confirmed that the energy efficiency was improved and can be further enhanced by improving the process efficiency.

The main contributions of this study are as follows. The dyeing process was parameterized based on research and advanced automatic configuration using several machine learning techniques. Inaccurate process instruction from the experiment in the laboratory were replaced by the CPES utilizing manufacturing big data, and invalid and ineffective steps in the traditional work process derived from operator's experience removed. The energy consumption of the dyeing process can be reduced by utilizing the CPES instead of purchasing expensive machines. As a result of application of the CPES, it was found that the energy consumption decreased by approximately 10.69% compared to the existing dyeing process. In addition, field operators will be able to cope with the possibility of repeated-dyeing (15–20% in the existing process) in advance through the CPES.

In the future, when further manufacturing big data are accumulated and sufficient data are available, the accuracy of the model could be improved by applying ensemble learning. We plan to extend the knowledge gained in this research to entire manufacturing processes in dyeing and finishing shops. In addition, we plan to improve the maturity level of the CPES to a self-optimizing level through application of an automatic dyeing machine controller.

References

1. Park, C.-W., Kwon, K.-S., Kim, W.-B., Min, B.-K., Park, S.-J., Sung, I.-H., et al. (2009). Energy consumption reduction technology in manufacturing—a selective review of policies, standards, and research. *International Journal of Precision Engineering and Manufacturing, 10*(5), 151–173.
2. Dornfeld, D. A. (2014). Moving towards green and sustainable manufacturing. *International Journal of Precision Engineering and Manufacturing-Green Technology, 1*(1), 63–66.
3. Hermann, C., Schmidt, C., Kurle, D., Blume, S., & Thiede, S. (2014). Sustainability in manufacturing and factories of the future. *International Journal of Precision Engineering and Manufacturing-Green Technology, 1*(4), 283–292.
4. Müller, J. M., & Kai-Ingo, V. (2018). Sustainable industrial value creation in SMEs: A comparison between industry 4.0 and made in China 2025. *International Journal of Precision Engineering and Manufacturing-Green Technology, 5*(5), 659–670.
5. Zhao, W.-B., Jeong, J.-W., Noh, S.-D., & Yee, J. T. (2015). Energy simulation framework integrated with green manufacturing-enabled PLM information model. *International Journal of Precision Engineering and Manufacturing-Green Technology, 2*(3), 217–224.
6. Beier, G., Niehoff, S., Ziems, T., & Xue, B. (2017). Sustainability aspects of a digitalized industry—a comparative study from China and Germany. *International Journal of Precision Engineering and Manufacturing-Green Technology, 4*(2), 227–234.
7. Palensky, P., Widl, E., & Elsheikh, A. (2013). Simulating cyber-physical energy systems: Challenges, tools and methods. *IEEE Transactions on Systems, Man, and Cybernetics, 44*(3), 318–326.
8. Kang, H.-S., Lee, J.-Y., Choi, S.-S., Kim, H., Park, J.-H., Son, J. Y., et al. (2016). Smart manufacturing: Past research, present findings, and future directions. *International Journal of Precision Engineering and Manufacturing-Green Technology, 3*(1), 111–128.
9. Elsheikh, A., Awais, M. U., Widl, E., Palensky, P. (2013) Modelica-enabled rapid prototyping of cyber-physical energy systems via the functional mockup interface. In 2013 workshop on modeling and simulation of cyber-physical energy systems (MSCPES).
10. KOSIS, Energy consumption by industry sector. Retrieved 10 September, 2017 from http://kosis.kr/statHtml/statHtml.do?orgId=337&tblId=TX_337_2013_H1017&vw_cd=MT_ZTITLE&list_id=337_33703_002&seqNo=&lang_mode=ko&language=kor&obj_var_id=&itm_id=&conn_path=E1#.
11. Korea Chemical Fibers Association (2013) Energy efficiency improvement guide. Development of energy efficiency improvement guide for textile industry research Report.
12. Park, K.-T., Yang, S.-G., Park, H.-J., Zhao, W.-B., Noh, S.-D., et al. (2017). A study on utilization of manufacturing big data for energy efficiency of dyeing-finishing industry. *Korean Journal of Computational Design and Engineering, 22*(4), 435–444.

13. Park, H.-S., & Nguyen, T. T. (2014). Optimization of injection molding process for car fender in consideration of energy efficiency and product quality. *Journal of Computational Design and Engineering, 1*(4), 256–265.
14. Monostori, L., Kádár, B., Bauernansl, T., Kondoh, S., Kumara, S., Reinhart, G., et al. (2016). Cyber-physical systems in manufacturing. *CIRP Annals, 65*(2), 621–641.
15. Gabor, T., Belzner, L., Kiermeier, M., Beck, M. T., Neitz, A. (2016) A simulation-based architecture for smart cyber-physical systems. In 2016 IEEE international conference on autonomic computing (ICAC), Würzburg, Germany, pp. 374–379.
16. Llic, M. D., Xie, L., Khan, U. A., & Moura, J. M. (2010). Modeling of future cyber-physical energy systems for distributed sensing and control. *Transactions on Systems, Man, and Cybernetics, 40*(4), 825–838.
17. Korea Institute of Industrial Technology (2016) Expansion of dyeing expert system (1/2). Korea Institute of Industrial Technology Research Report.
18. Korea Institute for Industrial Economics & Trade (2002) A study on the status analysis and establishment of process improvement direction for energy saving in dyeing. Korean Institute for Industrial Economics & Trade Research Report.
19. Korea Institute of Industrial Technology (1997) Development of dye/preparations to reduce re-infiltration rate of dyeing finishing factory. Korea Institute of Industrial Technology Research Report.
20. Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: Synthetic minority over-sampling technique. *Journal of Artificial Intelligence Research, 16,* 321–357.
21. Kohonen, T. (1988). An introduction to neural computing. *Neural Network, 1*(1), 3–16.
22. Barschdorff, D., & Monostori, L. (1991). Neural networks: Their applications and perspectives in intelligent machining. *Computers in Industry, 17*(2–3), 101–119.
23. Huang, S. H., & Zhang, H.-C. (1994). Artificial neural networks in manufacturing: Concepts, applications, and perspectives. *IEEE Transactions on Components, Packaging, and Manufacturing Technology, Part A, 17*(2), 212–228.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Kyu Tae Park 2013–2017: Bachelor of Systems Management Engineering, Sungkyunkwan University. 2016–current: Master and Ph.D. Integrated Course of Industrial Engineering, Sungkyunkwan University. Research interest: Cyber physical system, Interoperability, Smart manufacturing, Sustainable manufacturing, Service-oriented architecture.


Yong Tae Kang 1985–1990: Bachelor of Industrial Engineering, Hanyang University. 1991–1992: Master of Applied Statistics, KAIST. 1992–2006: Daewoo Information System Automation Division/SFC Team Leader. 2006-current: Head director of Research Institute in CoEver I & T Co., Ltd. Research interest: ICT based Factory/Building Automation, Big data-based Automation Solution, Cyber Physical System.

Suk Gon Yang, Ph.D. 1995–2001: Bachelor of Textile Engineering, Yeungnam University. 2009–2011: Master of Business Administration, Kyungpook National University. 2013–2018: Ph.D. of Business Administration, Yeungnam University. 2001-current: Center Chief of Advanced Processing Developing Center. Research interest: Energy System, Production Automation System, Smart Factory.

Wen Bin Zhao, Ph.D. 2005–2009: Bachelor of Industrial Engineering, Kangwon National University. 2009–2011: Master of Industrial Engineering, Sungkyunkwan University. 2011–2017: Ph.D. of Industrial Engineering, Sungkyunkwan University. 2017-current: Associated Professor of Taiyuan University of Technology. Research interest: CAD/CAPP/CAM, Digital Virtual Manufacturing, PLM, Engineering Interoperability, Smart Manufacturing, Smart Factory, Sustainable Manufacturing.
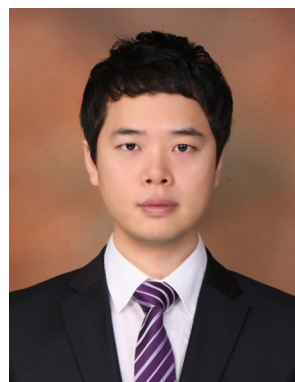
Yong-Shin Kang, Ph. D. 1999–2004: Bachelor of Industrial Systems Engineering, Dongguk University. 2004–2006: Master of Industrial Systems Engineering, Dongguk University. 2006–2011: Ph.D. of Industrial Systems Engineering, Dongguk University. 2013–2016: Associate Professor, Department of Industrial Systems Engineering,

Dongguk University. 2016-current: Research Professor, Department of Systems Management Engineering, Sungkyunkwan University. Research interest: Big data, IoT, CPS, SOA, traceability, trajectory.

Sung Ju Im 2011–2018: Bachelor of Systems Management Engineering, Sungkyunkwan University. 2017–current: Master Course of Industrial Engineering, Sungkyunkwan University. Research interest: CPPS, Industrial Automation, Smart Manufacturing.

Dong Hyun Kim 2002–2007: Bachelor of Computer Science, Chungju University. 2007-current: CoEver I & T Co., Ltd. Research interest: S/W development, IoT, smart factory, Big Data.

Su Young Choi 2002–2007: Bachelor of Computer Science, Chungju University. 2007-current: CoEver I & T Co., Ltd. Research interest: S/W development, machine running, IoT, smart manufacturing, smart factory, Big Data.

Sang Do Noh, Ph.D. 1988–1992: B.Sc., Mechanical Engineering, KAIST, Korea. 1992–1999: M.Sc/Ph.D., Manufacturing Engineering, Seoul National University, Korea. 1999–2002: Senior Research Engineer, Institute for Advanced Engineering. 2002-current: Professor, Department of Systems Management Engineering/Department of Industrial Engineering, Sungkyunkwan University. 2008–2009: Visiting Researcher, GM Global R&D Center. 2017–2018: Guest Professor, Royal Institute of Technology/Mälardalen University. Research interest: CAD/CAPP/CAM, Digital Virtual Manufacturing, PLM, Smart Manufacturing, Smart Factory, CPPS, Digital Twin.