DESCRIPTION

Title of Invention: COMPUTER PROGRAM, LEARNING MODEL GENERATION METHOD, INFORMATION PROCESSING METHOD, AND INFORMATION PROCESSING APPARATUS

Technical Field

[0001]

The present disclosure relates to a computer program, a learning model generation method, an information processing method, and an information processing apparatus.

Background Art

[0002]

PTL 1 proposes an anomaly detection apparatus in which an encoder compresses input data using a compression parameter fitted to normal data, a first anomaly in the input data is detected based on a log probability density of compressed data on a prior distribution and a log density ratio between a compressed data distribution and the prior distribution, a decoder decodes the compressed data using a decoding parameter fitted to the normal data, and a second anomaly in the input data is detected based on a difference between decoded reconstructed data and the input data.

Citation List

Patent Literature

[0003]

PTL 1: JP2018-49355A

Summary of Invention

Technical Problem

[0004]

The present disclosure provides a computer program, a learning model generation method, an information processing method, and an information processing apparatus that can be expected to reduce an influence of noise or the like in substrate processing state detection using a learning model.

Solution to Problem

[0005]

A computer program according to an embodiment is a computer program for causing a

computer to execute processing of generating, by machine learning, a learning model that receives, as an input, sensor data including a plurality of measurement values measured by a plurality of sensors related to substrate processing, and outputs sensor data including a plurality of corresponding measurement values, in which the computer program causes the computer to execute processing of acquiring the sensor data related to the substrate processing, inputting the acquired sensor data into the learning model and acquiring the sensor data output by the learning model, calculating an individual error for each measurement value between the sensor data input into the learning model and the sensor data output by the learning model, assigning weights to a plurality of the calculated individual errors, respectively, to calculate a total error, updating an internal parameter of the learning model based on the calculated total error, and adjusting the weights assigned to the individual errors according to a convergence status of the individual errors.

Advantageous Effects of Invention

[0006]

According to the present disclosure, it can be expected that an influence of noise or the like is reduced in substrate processing state detection using a learning model.

Brief Description of Drawings

[0007]

[FIG.1] FIG.1 is a schematic diagram illustrating an overview of an information processing system according to the present embodiment.

[FIG.2] FIG.2 is a block diagram illustrating a configuration example of an information processing apparatus according to the present embodiment.

[FIG.3] FIG.3 is a schematic diagram illustrating a configuration example of a learning model according to the present embodiment.

[FIG.4] FIG.4 is a flowchart illustrating an example of a procedure of learning model generation processing performed by the information processing apparatus according to the embodiment.

[FIG.5] FIG.5 is a flowchart illustrating an example of a procedure of anomaly detection processing performed by the information processing apparatus according to the embodiment.

[FIG.6] FIG.6 is a schematic diagram illustrating a display example of an anomaly detection result.

[FIG.7] FIG.7 is a schematic diagram illustrating data augmentation processing performed by the information processing apparatus according to the embodiment.

[FIG.8] FIG.8 is a schematic diagram illustrating preprocessing performed by the

information processing apparatus according to the embodiment.

Description of Embodiments

[0008]

Hereinafter, a specific example of an information processing system according to the embodiment of the present disclosure will be described with reference to the drawings. The present disclosure is not limited to these examples, and is defined by the claims, and is intended to include all modifications within the meaning and scope equivalent to the claims.

[0009]

<System Overview>

FIG. 1 is a schematic diagram illustrating an overview of an information processing system according to the present embodiment. The information processing system according to the present embodiment includes an information processing apparatus 1 and a substrate processing apparatus 3. The substrate processing apparatus 3 is an apparatus that performs various types of processing such as chemical vapor deposition (CVD), sputtering, or etching on a substrate (wafer) of a semiconductor, for example. The substrate processing apparatus 3 according to the embodiment includes a plurality of sensors 4, such as a sensor for measuring temperature and a sensor for measuring pressure in a chamber accommodating the substrate to be processed. The substrate processing apparatus 3 repeatedly acquires (samples), at a predetermined interval, data such as the temperature and the pressure measured by the plurality of sensors 4 when performing substrate processing such as etching, and transmits the data acquired from the plurality of sensors 4 collectively to the information processing apparatus 1 as sensor data. Alternatively, the substrate processing apparatus 3 may store, in a memory or the like, the data acquired from the plurality of sensors 4 when performing the substrate processing and transmit the stored data collectively to the information processing apparatus 1 after completion of the substrate processing. Information exchange between the substrate processing apparatus 3 and the information processing apparatus 1 may be performed via a recording medium such as a memory card or an optical disc, instead of wired or wireless communication.

[0010]

The information processing apparatus 1 is an apparatus that performs processing of detecting an anomaly or the like related to substrate processing performed by the substrate processing apparatus 3, based on the sensor data obtained from the plurality of sensors 4 provided in the substrate processing apparatus 3, for example. There may be a plurality of substrate processing apparatuses 3 whose anomaly or the like is to be detected by the information processing apparatus 1. The information processing apparatus 1 according to the

embodiment collects the sensor data of the substrate processing apparatus 3 and records the sensor data in a database, and generates a learning model 2 for detecting the anomaly or the like by performing machine learning using the recorded sensor data.

[0011]

The learning model 2 according to the embodiment is a learning model adopting a configuration of a so-called autoencoder. The autoencoder is a learning model that combines an encoder that compresses received data to transform the data into feature data with reduced dimensionality, and a decoder that restores the feature data to data with original dimensionality. Machine learning of the autoencoder is performed by a so-called unsupervised machine learning method that does not require training data in which a ground truth value is assigned to input data. The information processing apparatus 1 can perform the machine learning on the autoencoder learning model 2 by updating an internal parameter such that input data and output data match each other.

[0012]

When performing anomaly detection using the autoencoder, normal-time data is used for the machine learning of the autoencoder and abnormal-time data is not required. The autoencoder subjected to the machine learning using the normal-time data outputs normal-time data when the normal-time data is received, and does not output abnormal-time data even when the abnormal-time data is received. The information processing apparatus 1 can input, into the autoencoder learning model 2, data whose normal or abnormal state is unknown, acquire data output by the learning model 2, and determine that the received data has an anomaly when a difference between the input data and the output data exceeds a predetermined threshold value. A method for training the autoencoder learning model and a method for detecting an anomaly using the autoencoder learning model are techniques known in the related art, and thus detailed descriptions thereof are omitted.

[0013]

The information processing apparatus 1 according to the embodiment records the sensor data obtained from the substrate processing apparatus 3 in the database, and performs the machine learning of the autoencoder learning model 2 using the sensor data when the substrate processing is normally performed. Thereafter, the information processing apparatus 1 detects, using the learning model 2 trained by the machine learning, an anomaly in the substrate processing performed by the substrate processing apparatus 3. The information processing apparatus 1 acquires the sensor data measured by the plurality of sensors 4 in association with the substrate processing from the substrate processing apparatus 3, inputs the acquired sensor data to the learning model 2, and calculates a difference between the input sensor data and data output by the learning model 2. When the calculated difference exceeds a predetermined

threshold value, the information processing apparatus 1 determines that there is an anomaly in the substrate processing and notifies a user such as an administrator of this information processing system of the anomaly.

[0014]

In the information processing system that performs anomaly detection using such an autoencoder learning model 2, an electrical signal output as a measurement result of the temperature or the pressure by the plurality of sensors 4 provided in the substrate processing apparatus 3 may include noise unrelated to the substrate processing. In such a case, the information processing apparatus 1 may erroneously detect the noise in the acquired sensor data as an anomaly. In the information processing system according to the embodiment, the machine learning is performed on the learning model 2 to prevent erroneous detection due to the noise in the sensor data.

[0015]

<Apparatus Configuration>

FIG. 2 is a block diagram illustrating a configuration example of the information processing apparatus 1 according to the present embodiment. The information processing apparatus 1 according to the present embodiment can be implemented by installing a given application program or the like in a general-purpose information processing apparatus such as a personal computer or a server computer. The information processing apparatus 1 according to the present embodiment includes a processor 11, a storage 12, a communication unit 13, a display unit 14, an operation unit 15, and the like. In the present embodiment, an example will be described in which a process is performed by one information processing apparatus 1. Meanwhile, the process of the information processing apparatus 1 may be distributed and performed by a plurality of apparatuses.

[0016]

The processor 11 is configured by using an arithmetic processing apparatus such as a central processing unit (CPU), a micro-processing unit (MPU), a graphics processing unit (GPU), or a quantum processor, a read only memory (ROM), a random access memory (RAM), and the like. The processor 11 reads and executes a program 12a stored in the storage 12, thereby performing processing of generating the learning model 2 based on the sensor data obtained from the sensors 4 of the substrate processing apparatus 3, various types of processing such as substrate processing anomaly detection using the generated learning model 2.

[0017]

The storage 12 is configured by using, for example, a large-capacity storage device such as a hard disk or a solid state drive (SSD). The storage 12 stores various types of programs to be executed by the processor 11 and various types of data necessary for the process of the

processor 11. In the present embodiment, the storage 12 stores the program 12a to be executed by the processor 11. The storage 12 includes a sensor data storage 12b for storing the sensor data acquired from the substrate processing apparatus 3 and a model information storage 12c for storing information on the learning model 2 generated by the information processing apparatus 1.

[0018]

In the present embodiment, the program (computer program, program product) 12a is provided in a form recorded on a recording medium 99 such as a memory card or an optical disc. The information processing apparatus 1 reads the program 12a from the recording medium 99, and stores the program 12a in the storage 12. However, for example, the program 12a may be written into the storage 12 during a manufacturing stage of the information processing apparatus 1. For example, as the program 12a, the information processing apparatus 1 may acquire those which are distributed by a remote server device or the like through communication. For example, the program 12a may be written into the storage 12 of the information processing apparatus 1 after a writing apparatus reads data recorded in the recording medium 99. The program 12a may be provided in the form of distribution through a network, or may be provided in the form recorded in the recording medium 99.

[0019]

The sensor data storage 12b stores the sensor data acquired from the substrate processing apparatus 3 together with information such as date and time when the sensor data is acquired, identification information of the substrate processing apparatus 3, and identification information of a substrate processed by the substrate processing apparatus 3. The substrate processing apparatus 3 is provided with the plurality of sensors 4, and the sensor data is obtained by repeatedly sampling signals output by the plurality of sensors 4 at a predetermined interval when performing the substrate processing. For example, when the substrate processing apparatus 3 includes M sensors and N times of sampling are performed during the substrate processing, the sensor data acquired by the information processing apparatus 1 includes $M \times N$ values. The sensor data storage 12b stores the sensor data including the $M \times N$ values in association with information such as date and time and identification information.

[0020]

The model information storage 12c stores information on the learning model subjected to the machine learning. The information related to the learning model may include, for example, information indicating a configuration of the learning model, and information such as values of internal parameters determined by machine learning. In the embodiment, the model information storage 12c stores the information on the autoencoder learning model 2 used for anomaly detection for the substrate processing apparatus 3. The learning model 2 according to

the embodiment is a learning model subjected to machine learning to receive, as an input, the sensor data including the M × N values and outputs a normal sensor data prediction value for such sensor data as M × N pieces of data. The information processing apparatus 1 performs the machine learning using the normal sensor data acquired from the substrate processing apparatus 3 to generate the learning model 2, and stores the information on the generated learning model 2 in the model information storage 12c.

[0021]

In the present embodiment, the information processing apparatus 1 stores the information related to the learning model, and the information processing apparatus 1 executes processing based on the learning model. However, the present disclosure is not limited thereto. The information related to the learning model may be stored in a device different from the information processing apparatus 1, the device may execute processing based on the learning model, and the information processing apparatus 1 may acquire a processing result from the device. Further, machine learning processing of the learning model may be executed by the information processing apparatus 1, or may be executed by a device different from the information processing apparatus 1.

[0022]

The communication unit 13 is connected to the substrate processing apparatus 3 through a communication line, for example, and exchanges data with the substrate processing apparatus 3. In the embodiment, the communication unit 13 receives the sensor data transmitted from the substrate processing apparatus 3 and supplies the sensor data to the processor 11. The communication unit 13 transmits data such as a control command from the processor 11 to the substrate processing apparatus 3.

[0023]

The display unit 14 is configured by using a liquid crystal display or the like, and displays various images, characters, and the like based on the process of the processor 11. The display unit 14 displays various types of information, such as the sensor data acquired from the substrate processing apparatus 3, the information on the generated learning model 2, and an operation status of the controlled substrate processing apparatus 3.

[0024]

The operation unit 15 receives a user operation and notifies the processor 11 of the received operation. For example, the operation unit 15 receives the user operation by an input device such as a mechanical button or a touch panel provided on a surface of the display unit 14. For example, the operation unit 15 may be an input device such as a mouse and a keyboard, and these input devices may be configured to be detachable from the information processing apparatus 1.

[0025]

The storage 12 may be an external storage device connected to the information processing apparatus 1. The information processing apparatus 1 may be a multi-computer including a plurality of computers, or may be a virtual machine virtually constructed by software. In addition, the information processing apparatus 1 is not limited to the configuration described above, and does not need to include the display unit 14, the operation unit 15, and the like, for example.

[0026]

In the information processing apparatus 1 according to the embodiment, the processor 11 reads and executes the program 12a stored in the storage 12 to implement a sensor data acquisition unit 11a, a learning processor 11b, an anomaly detector 11c, a display processor 11d, and the like in the processor 11 as software functional units. In this figure, functional units related to processing such as collecting the sensor data from the substrate processing apparatus 3, generating the learning model 2 and detecting an anomaly in the substrate processing apparatus 3 are illustrated as the functional units of the processor 11, and functional units related to other processing are not illustrated.

[0027]

By communicating with the substrate processing apparatus 3 through the communication unit 13, the sensor data acquisition unit 11a performs processing of acquiring the sensor data including a plurality of measurement values such as the temperature and the pressure measured by the plurality of sensors 4 when the substrate processing apparatus 3 performs the substrate processing. The substrate processing apparatus 3 samples the signals output by the plurality of sensors 4 at a predetermined interval at the time of the substrate processing to store the sensor data in a memory or the like, and transmits the stored sensor data to the information processing apparatus 1 at an appropriate timing after the substrate processing is ended, for example. The sensor data acquisition unit 11a of the information processing apparatus 1 acquires the sensor data transmitted from the substrate processing apparatus 3 after the substrate processing is ended, and stores the acquired sensor data in the sensor data storage 12b in association with information such as acquisition date and time and identification information on the substrate processing apparatus 3.

[0028]

Alternatively, the sensor data acquisition unit 11a may acquire the sensor data repeatedly by communicating with the substrate processing apparatus 3 at a predetermined interval when the substrate processing apparatus 3 performs the substrate processing. Accordingly, the sensor data acquisition unit 11a can acquire the sensor data in real time while the substrate processing is performed by the substrate processing apparatus 3. The interval and

a timing when the sensor data acquisition unit 11a acquires the sensor data from the substrate processing apparatus 3 may be appropriately determined by a designer of this information processing system.

[0029]

The learning processor 11b performs machine learning processing for generating the learning model 2 using the sensor data stored in the sensor data storage 12b. As described above, the learning model 2 has an autoencoder configuration, receives the sensor data including the M × N values as the input, and outputs the data including the M × N values similarly. The learning processor 11b uses sensor data when the substrate processing is normally performed among the sensor data stored in the sensor data storage 12b, and updates the internal parameter of the learning model 2 such that the sensor data input into the learning model 2 matches the data output by the learning model 2. Accordingly, the learning processor 11b generates the learning model 2 that predicts normal sensor data with respect to the input of the sensor data. The learning processor 11b stores, in the model information storage 12c, information such as the internal parameter of the learning model 2 obtained by the machine learning.

[0030]

In the embodiment, among the plurality of sensors 4 in the substrate processing apparatus 3, a sensor 4 for which prediction by the learning model 2 is difficult is a sensor whose sensor data includes noise unrelated to the substrate processing and is regarded as having a relatively small influence on subsequent anomaly detection and the like. In the machine learning for generating the learning model 2, the learning processor 11b calculates a prediction error for each sensor 4 and proceeds with the machine learning such that, by reducing learning importance for a sensor 4 whose error is difficult to converge, the learning model 2 is prevented from outputting sensor data including noise as correct sensor data. In addition, the learning processor 11b stores, in the model information storage 12c, information indicating which sensor 4 is determined to be difficult for prediction and has the learning importance thereof reduced, together with information such as the internal parameter of the learning model 2.

[0031]

The anomaly detector 11c performs processing of detecting, using the learning model 2 generated by the machine learning in advance, the anomaly related to the substrate processing based on the sensor data acquired from the substrate processing apparatus 3. The anomaly detector 11c reconstructs the learning model 2 based on the information stored in the model information storage 12c. The anomaly detector 11c acquires the sensor data measured by the plurality of sensors 4, for example, when the substrate processing apparatus 3 performs the substrate processing or after the substrate processing is ended. The anomaly detector 11c inputs the sensor data acquired from the substrate processing apparatus 3 into the learning model 2,

and acquires the data output by the learning model 2 in response thereto. The anomaly detector 11c determines whether there is the anomaly related to the substrate processing based on a difference between the data input into the learning model 2 and the data output by the learning model 2.

[0032]

The anomaly detector 11c may individually detect an anomaly for the plurality of sensors 4 provided in the substrate processing apparatus 3. For example, the anomaly detector 11c can calculate, for each sensor 4, an error such as a mean squared error with respect to the input data to the learning model 2 and the output data from the learning model 2, and determine that there is an anomaly in a substrate processing event measured by a sensor 4 whose calculated error exceeds a threshold value. At this time, the anomaly detector 11c may exclude the sensor 4, which is determined to be difficult for prediction and for which the learning importance is reduced in the machine learning of the learning model 2, from the anomaly detection and do not perform processing such as the error calculation and the threshold value comparison.

[0033]

The display processor 11d performs processing of displaying various characters and images on the display unit 14. In the embodiment, the display processor 11d converts the sensor data acquired from the substrate processing apparatus 3 into a graph and displays the graph on the display unit 14. The display processor 11d displays, on the display unit 14, a progress display of the machine learning processing performed by the learning processor 11b, for example, a graph illustrating a change in the prediction error of the learning model 2. In addition, the display processor 11d displays, on the display unit 14, information on a result of the substrate processing anomaly detection using the learning model 2, such as whether an anomaly is present in the substrate processing and, when the anomaly is present, information indicating which sensor 4 detects the anomaly. The display processor 11d may display various types of information other than the information described above on the display unit 14.

[0034]

<Learning Model Generation Processing>

FIG. 3 is a schematic diagram illustrating a configuration example of the learning model 2 according to the embodiment. The learning model 2 according to the embodiment is a learning model having an autoencoder configuration including an encoder and a decoder (not illustrated). The encoder and the decoder of the learning model 2 may adopt, for example, a neural network configuration. The encoder of the learning model 2 transforms received data into a feature having reduced dimensionality, and the decoder transforms the feature inversely into data having original dimensionality. The learning model 2 receives, as an input, the sensor data measured by the plurality of sensors 4 provided in the substrate processing apparatus 3,

and is subjected to the machine learning such that a predicted value for normal sensor data is output with respect to the received sensor data.

[0035]

The information processing apparatus 1 acquires the sensor data when the substrate processing apparatus 3 performs the substrate processing, and stores the acquired normal sensor data in the sensor data storage 12b. To the sensor data stored in the sensor data storage 12b, information indicating whether the substrate processing is normally performed, for example, is assigned by the user who checks a result of the substrate processing. In the embodiment, it is assumed that whether the sensor data is normal is determined by the user, but the determination is not limited thereto. For example, the information processing apparatus 1 may perform processing of measuring a resultant product of the substrate processing by a measuring instrument and determining whether a measurement value obtained by the measurement is within a predetermined range. Based on such a determination result, the information processing apparatus 1 may assign information indicating normal or abnormal to the sensor data and store the sensor data in the sensor data storage 12b. The information processing apparatus 1 can extract normal sensor data from the sensor data stored in the sensor data storage 12b based on such information. Using the sensor data when the substrate processing is normally performed among the sensor data stored in the sensor data storage 12b, the information processing apparatus 1 performs the machine learning processing for generating the learning model 2.

[0036]

In this figure, the sensor data serving as the input to the learning model 2 and the output therefrom is illustrated as a waveform. However, in practice, for example, when the substrate processing apparatus 3 includes M sensors and sampling is performed N times during the substrate processing, data input to and output from the learning model 2 can be represented by a matrix of M × N. The information processing apparatus 1 reads the sensor data stored in the sensor data storage 12b, and uses the read sensor data subjected to preprocessing such as normalization or standardization if necessary as training data for performing the machine learning of the learning model 2.

[0037]

The information processing apparatus 1 inputs the training sensor data read from the sensor data storage 12b to the learning model 2 where an appropriate initial value is set as the internal parameter in advance, and acquires the data output by the learning model 2. As described above, the sensor data input to the learning model 2 is data of an M × N matrix and the data output by the learning model 2 is also data of an M × N matrix. The information processing apparatus 1 according to the embodiment calculates a squared error for each element in the M × N matrix as an individual error $L^{t}_{current}$ using the following formula (1), based on the

input data to the learning model 2 and the output data from the learning model 2. By such calculation, the information processing apparatus 1 obtains M × N individual errors. In formula (1), x^t_{input} is an element of the matrix of the input data, and x^t_{pred} is an element of the matrix of the corresponding output data. In addition, a superscript t in formula (1) and subsequent formulas indicates the number of iterations, that is, the so-called number of epochs in the machine learning.

[0038]

[Formula 1]

$$\mathcal{L}^{t}_{\text{current}} = \left(x^{t}_{input} - x^{t}_{pred}\right)^2 \qquad \cdots (1)$$

[0039]

Next, the information processing apparatus 1 calculates an exponential moving average error L^t_{ema} for each individual error $L^t_{current}$ based on the following formula (2). By such calculation, the information processing apparatus 1 obtains M × N exponential moving average errors L^t_{ema}. In formula (2) and subsequent formulas, a coefficient α is a hyperparameter determined by the designer or the like of the information processing system according to the embodiment. The information processing apparatus 1 calculates a current exponential moving average error based on a current individual error and a previous exponential moving average error in an iteration of the machine learning. Therefore, the information processing apparatus 1 stores a calculation result of the current exponential moving average error in the storage 12 at least until next exponential moving average error calculation is performed.

[0040]

[Formula 2]

$$\mathcal{L}^{t}_{\text{ema}} = \mathcal{L}^{t}_{\text{current}} + \alpha\mathcal{L}^{t-1}_{ema} \qquad \cdots (2)$$

[0041]

Next, the information processing apparatus 1 calculates a machine learning weight coef. for each element in the M × N matrix of the sensor data based on the following formula (3). By such calculation, the information processing apparatus 1 obtains M × N weights coef. A coefficient β in (3) is a hyperparameter determined by the designer or the like of the information processing system according to the embodiment. For an element of the sensor 4 whose error is difficult to converge by machine learning iterations, a value of the weight coef. is reduced. The information processing apparatus 1 can assign weights such that, as an individual error at a certain time point becomes smaller than a previous individual error, a

weight assigned to this individual error becomes larger.

[0042]

[Formula 3]

$$\text{coef.} = \left(\frac{(1-\alpha)\mathcal{L}_{\text{ema}}^{t-1}}{\mathcal{L}_{\text{current}}^{t} + (1-\alpha)\mathcal{L}_{\text{ema}}^{t-1}} \right)^{\beta} \quad \cdots (3)$$

[0043]

Next, the information processing apparatus 1 calculates a total error L^t based on the following formula (4), using M × N individual errors $L^t_{current}$ calculated based on formulas (1) to (3) described above and M × N weights coef. corresponding thereto. The total error L^t is obtained by multiplying each of the M × N individual errors $L^t_{current}$ by a corresponding weight and calculating a sum of resulting M × N values, and is one numerical value.

[0044]

[Math 4]

$$\mathcal{L}^t = \sum coef. \times \mathcal{L}_{current}^{t} \quad \cdots (4)$$

[0045]

Using the total error L^t, the information processing apparatus 1 performs the machine learning of the learning model 2 by updating the internal parameter of the learning model 2 such that the calculated total error L^t decreases. The information processing apparatus 1 can update the internal parameter of the learning model 2, for example, using a backpropagation method. Processing of updating the internal parameter of the learning model 2 using the backpropagation method is a technique known in the related art, and thus a detailed description thereof will be omitted. The total error L^t calculated based on the formulas (1) to (4) described above is a value in which the weight coef. of a value related to the sensor 4 whose error is difficult to converge, such as a sensor 4 including large noise, is reduced, and thus importance, influence, or the like in the machine learning related to such a sensor 4 is reduced. By performing the machine learning based on such a total error L^t, the learning model 2 is not trained with a feature of, for example, a sensor 4 that has large noise and is difficult to predict, and is trained with a feature of, for example, a sensor 4 that has small noise and is easy to predict.

[0046]

FIG. 4 is a flowchart illustrating an example of a procedure of learning model generation processing performed by the information processing apparatus 1 according to the embodiment. The processor 11 of the information processing apparatus 1 according to the embodiment initializes a value of a variable t used in the present processing to one (step S1). The variable t is a variable for counting the number of machine learning iterations, and is implemented using, for example, a register in the processor 11 or a storage area in the storage 12. The learning processor 11b of the processor 11 acquires, as the sensor data to be used in the machine learning, the sensor data measured by the sensor 4 when the substrate processing is normally performed, among the sensor data stored in the sensor data storage 12b of the storage 12 (step S2).

[0047]

The learning processor 11b inputs the sensor data acquired in step S2 to the learning model 2 (step S3). The learning processor 11b acquires data output by the learning model 2 with respect to the sensor data received in step S3 (step S4). The learning processor 11b calculates, based on formula (1) described above, the individual error for each element in the sensor data input to the learning model 2 in step S3 and the output data from the learning model 2 acquired in step S4 (step S5). The learning processor 11b calculates, based on formulas (2) and (3) described above, the weight of each element in the sensor data, based on the individual error calculated in step S5 and the previous exponential moving average error (step S6). The learning processor 11b calculates, based on formula (4) described above, the total error based on the individual error calculated in step S5 and the weight calculated in step S6 (step S7). The learning processor 11b updates, based on the total error calculated in step S7, the internal parameter of the learning model 2 such that the total error is reduced, using a method such as the backpropagation method (step S8).

[0048]

The learning processor 11b determines whether to end the machine learning processing, based on a condition such as that the number of machine learning iterations reaches a predetermined number or that the total error calculated in step S7 is less than a predetermined target value (step S9). When the machine learning processing is not ended (S9: NO), the learning processor 11b adds 1 to the variable t (step S10), and returns the processing to step S3. When the machine learning is ended (S9: YES), the learning processor 11b stores information such as the internal parameter of the final learning model 2 and the weight of each element in the sensor data obtained in a process of the machine learning, into the model information storage 12c (step S11), and ends the processing. The weight of each element stored in step S11 may be a representative value, such as a minimum value or an average value among a plurality of weights obtained by machine learning iterations.

[0049]

<Anomaly Detection Processing>

The information processing apparatus 1 according to the embodiment performs processing of detecting, using the learning model 2 generated by the machine learning described above, the anomaly related to the substrate processing of the substrate processing apparatus 3. The information processing apparatus 1 according to the embodiment acquires the sensor data including various measurement values measured, for example, when the substrate processing apparatus 3 performs the substrate processing or after the substrate processing is ended, by the plurality of sensors 4 in the process of the substrate processing. The information processing apparatus 1 inputs the acquired sensor data (or data obtained by performing appropriate preprocessing on the sensor data) into the trained learning model 2, and acquires the data output by the learning model 2 in response thereto.

[0050]

The information processing apparatus 1 compares each element of the data input to the learning model 2 with that output by the learning model 2, and determines, when a difference (error) of each element exceeds a predetermined threshold value, that there is an anomaly in the sensor 4 related to the element whose difference exceeds the threshold value. The threshold value used for the anomaly detection may be a common value with respect to a plurality of elements, or an individual value may be set for each element, for example. These threshold values are determined in advance by the designer or the like of the information processing system according to the embodiment, and are stored in advance in the storage 12 of the information processing apparatus 1. The difference of each element may be, for example, a difference value or an absolute value of two values of input and output, for example, the individual error obtained by formula (1) described above may be used, or a value obtained by other calculation may be used.

[0051]

The information processing apparatus 1 may determine whether an anomaly is present for each sensor 4 instead of determining whether an anomaly is present for each element of input and output data. For example, when the sampling is performed N times for M sensors 4 to obtain the sensor data, the input and output data with respect to the learning model 2 is a matrix of M × N. When determining whether an anomaly is present for each element, the information processing apparatus 1 may calculate M × N differences and compare the differences with the threshold value. When determining whether an anomaly is present for each sensor 4, the information processing apparatus 1 can treat N values for each sensor 4 as an N-dimensional vector, calculate a similarity between two N-dimensional vectors of input and output, and determine that a sensor 4 having a similarity less than a threshold value is abnormal.

[0052]

As described above, in the information processing apparatus 1 according to the embodiment, at the time of the machine learning for generating the learning model 2, the weight is set to be small for the sensor 4 which includes large noise and whose error is difficult to converge, and the feature thereof is not used for training. Therefore, the trained learning model 2 has low prediction accuracy for the sensor 4 whose weight is set to be small and may not accurately detect an anomaly. The information processing apparatus 1 stores the weight calculated at the time of the machine learning, together with the information such as the internal parameter of the learning model 2 obtained by the machine learning, and excludes the sensor 4 including an element whose weight is lower than a predetermined threshold value from the anomaly detection. The information processing apparatus 1 may store a minimum value or an average value of the weight for each sensor 4, for example, instead of storing all weights obtained at the time of the machine learning, and may alternatively store identification information on the sensor 4 excluded from the anomaly detection determined based on the weight, instead of storing the weight.

[0053]

The information processing apparatus 1 displays, on the display unit 14, information on an anomaly detection result. The information processing apparatus 1 displays a list in which identification information on the plurality of sensors 4 provided in the substrate processing apparatus 3, for example, is associated with presence or absence of an anomaly in each sensor 4. At this time, the information processing apparatus 1 may display information indicating which sensor 4 is excluded from the anomaly detection. The information processing apparatus 1 may perform a more detailed information display, such as displaying the sensor data in the form of a graph for the sensor 4 where an anomaly is detected.

[0054]

FIG. 5 is a flowchart illustrating an example of a procedure of anomaly detection processing performed by the information processing apparatus 1 according to the embodiment. The sensor data acquisition unit 11a of the processor 11 of the information processing apparatus 1 according to the embodiment communicates with the substrate processing apparatus 3 through the communication unit 13 to acquire the sensor data including measurement values from the plurality of sensors 4 provided in the substrate processing apparatus 3 (step S31).

[0055]

The anomaly detector 11c of the processor 11 constructs the learning model 2 based on the information stored in the model information storage 12c, and inputs the sensor data acquired in step S31 into the learning model 2 (step S32). The anomaly detector 11c acquires data output by the learning model 2 with respect to the sensor data input in step S32 (step S33). The anomaly

detector 11c compares the sensor data input to the learning model 2 in step S32 with the output data of the learning model 2 acquired in step S33, and calculates an error of each element in the input and output data based on, for example, formula (1) described above (step S34).

[0056]

The anomaly detector 11c acquires the information on the weight of each element stored in the model information storage 12c when performing the machine learning of the learning model 2, and excludes the sensor 4 including the element whose weight is lower than the predetermined threshold value from the anomaly detection (step S35). The anomaly detector 11c determines whether the error calculated in step S34 exceeds the predetermined threshold value for each sensor 4 that is not excluded in step S35 (step S36). When the error exceeds the threshold value (S36: YES), the anomaly detector 11c notifies that an anomaly is detected in the sensor 4 whose error exceeds the threshold value (step S37), and ends the processing. When the error does not exceed the threshold value (S36: NO), the anomaly detector 11c ends the processing.

[0057]

FIG. 6 is a schematic diagram illustrating a display example of an anomaly detection result. The information processing apparatus 1 according to the embodiment displays, on the display unit 14, information on the sensor 4 targeted for the anomaly detection, in association with an anomaly degree (error) of each sensor, as a result of the substrate processing anomaly detection based on the sensor data acquired from the substrate processing apparatus 3. In the example illustrated in FIG. 6, based on the error of each sensor 4 obtained by the anomaly detection processing, the information processing apparatus 1 displays the plurality of sensors 4 provided in the substrate processing apparatus 3 in descending order of the anomaly degree (descending order of the error) and thus displays the information on the sensor 4 targeted for the anomaly detection in association with the anomaly degree. In this example, the information processing apparatus 1 displays the error of each sensor 4 as a bar graph in an upper portion of a screen and sequentially arranges the sensors 4 from left to right in descending order of error magnitude.

[0058]

The display of the anomaly detection result by the information processing apparatus 1 is not limited to the example illustrated in FIG. 6, and various display modes may be adopted. The information processing apparatus 1 may display the anomaly degree (error) of each sensor 4 using various graphs such as a line graph or a pie chart, instead of the bar graph. The information processing apparatus 1 may display the information on each sensor 4 in a tabular format instead of the graph, and in this case, for example, the information on the sensor 4 can be displayed in a list in the tabular format in descending order of the anomaly degree. The

information processing apparatus 1 may display the information on the sensor 4 not in order of the anomaly degree but, for example, in an appropriate order such as in order of the identification number of the sensor 4, and may highlight those with a high anomaly degree to display the information on the sensor 4 targeted for the anomaly detection in association with the anomaly degree. A method for highlighting may be, for example, changing a color, a size or font of characters for displaying the information according to the anomaly degree, or flashing the information on the sensor 4 having a high anomaly degree.

[0059]

The information processing apparatus 1 excludes the sensor 4 whose individual error weight is set to be low at the time of the machine learning from display targets, or performs display such that the sensor 4 is identifiable as being excluded from the anomaly detection. For example, the information processing apparatus 1 can display the information on the sensor 4 targeted for the anomaly detection and information on a non-target sensor 4 in a distinguishable manner by providing a difference in the display mode, such as by displaying the information on the target sensor 4 and the information on the non-target sensor 4 in different colors, sizes, or the like. For example, the information processing apparatus 1 may display the information on the non-target sensor 4 with an icon indicating that the information is excluded from the anomaly detection.

[0060]

In this figure, for a sensor 4 having a largest error (a sensor K in the figure), the information processing apparatus 1 displays a graph of time series sensor data used for the anomaly detection and a graph of time series errors calculated with respect to the time series sensor data, side by side at a lower portion in the screen. In this example, the information processing apparatus 1 displays the graphs of the sensor data and the errors for the sensor 4 having the largest error, but the display is not limited thereto. For example, the user can perform an operation of selecting a sensor name of any of a plurality of sensors 4 arranged at the upper portion in the screen in descending order of error, and the information processing apparatus 1 that receives this operation displays graphs of sensor data and errors related to the sensor 4 selected by the user at the lower portion in the screen.

[0061]

<Additional Function 1: Data Augmentation>

The information processing apparatus 1 performs machine learning to generate the learning model 2 using a plurality of pieces of sensor data acquired from the substrate processing apparatus 3 and stored in the sensor data storage 12b. As the training data used for the machine learning increases, prediction accuracy of the generated learning model 2 is expected to be improved. Therefore, the information processing apparatus 1 according to the

embodiment can perform processing of increasing the training data, that is, data augmentation by generating new sensor data that can be used for the machine learning based on the sensor data acquired from the substrate processing apparatus 3.

[0062]

FIG. 7 is a schematic diagram illustrating data augmentation processing performed by the information processing apparatus 1 according to the embodiment. The information processing apparatus 1 according to the embodiment uses, as original data, a time series measurement value of each sensor 4 in the sensor data acquired from the substrate processing apparatus 3 in advance, and performs appropriate calculation on the original data to generate one or a plurality of pieces of augmented data. The information processing apparatus 1 determines, as a range in which the data is to be augmented, a range obtained by augmenting the original data by a predetermined amount in each direction of a vertical axis (measurement value) and a horizontal axis (time). An upper portion in FIG. 7 illustrates a step waveform as an example of the original data, and a middle portion illustrates an example of an augmentation range for this waveform.

[0063]

In the embodiment, the information processing apparatus 1 determines the augmentation range for the original data based on information on a resolution (determined by a minimum value measurable by the sensor 4 for the vertical axis and by a sampling period for the horizontal axis, for example) of the sensor 4 that outputs the original data. The information processing apparatus 1 can set, as the augmentation range, a range in which the original data is translatable by plus or minus one resolution along each of the vertical axis and the horizontal axis, for example.

[0064]

The information processing apparatus 1 generates one or more pieces of augmented data by performing an operation such as translation, enlargement, reduction and the like on the original data such that resulting data falls within the determined augmentation range. Three augmented data waveforms generated to be within the augmentation range are illustrated at a lower portion in FIG. 7. Augmented data illustrated on a left side is obtained by reducing the original data in a vertical axis direction. Augmented data illustrated at a center is obtained by enlarging the original data in the vertical axis direction. Augmented data illustrated on a right side is obtained by translating the original data along the vertical axis direction. The information processing apparatus 1, for example, randomly selects a plurality of types of calculation such as predetermined translation, enlargement, or reduction, and performs the selected calculation on the original data to augment the data within the augmentation range.

[0065]

The information processing apparatus 1 can generate one or more pieces of augmented sensor data with respect to one piece of sensor data (normal sensor data) stored in the sensor data storage 12b, and thus increase an amount of the sensor data usable for the machine learning of the learning model 2 by a factor of two or more. By performing the machine learning using more training data, it is expected that the information processing apparatus 1 can generate the learning model 2 with higher prediction accuracy. In addition, by determining the augmentation range for performing the data augmentation based on the resolution of the sensor 4, it is expected that erroneous anomaly detection due to the resolution of the sensor 4 can be prevented.

[0066]

<Additional Function 2: Preprocessing>

The information processing apparatus 1 may perform appropriate preprocessing on the sensor data acquired from the substrate processing apparatus 3 and input the preprocessed sensor data to the learning model 2. For the preprocessing, for example, processing such as standardization or normalization may be adopted. The standardization is processing of transforming, based on an average value and a standard deviation of a plurality of values, the plurality of values such that the average value becomes 0 and a variance becomes 1. The normalization is processing of transforming, based on a minimum value and a maximum value of a plurality of values, the plurality of values into values within a range from 0 to 1. Processing such as the standardization and the normalization is a technique known in the related art, and thus a detailed description thereof will be omitted.

[0067]

FIG. 8 is a schematic diagram illustrating the preprocessing performed by the information processing apparatus 1 according to the embodiment. In the embodiment, the information processing apparatus 1 performs the preprocessing using a summary value of the sensor data and inputs the preprocessed sensor data into the learning model 2. The summary value of the sensor data is an average value and a standard deviation when the standardization is performed as the preprocessing, or a minimum value and a maximum value when the normalization is performed as the preprocessing. The information processing apparatus 1 may perform preprocessing different from the standardization and the normalization, and in this case, the summary value of the sensor data may be various values required for the preprocessing. Hereinafter, a case where the standardization is performed as the preprocessing will be described.

[0068]

The information processing apparatus 1 can obtain the sensor data including M × N values by performing sampling N times for M sensors 4, for example. The information

processing apparatus 1 can obtain two summary values by calculating an average value and a standard deviation of N values for each sensor 4, and can obtain 2 × M summary values from one piece of sensor data.

[0069]

The information processing apparatus 1 according to the embodiment includes a summary value prediction model 5 generated in advance by machine learning (that is, stores information on the prediction model 5 in the model information storage 12c). The summary value prediction model 5 is a learning model that is subjected to the machine learning to receive, as an input, the summary value, and predict a normal summary value corresponding thereto. In the embodiment, the information processing apparatus 1 uses, as the summary value prediction model 5, a model that performs dimensionality reduction by principal component analysis, transforms 2 × M summary values into information of a smaller number of principal components, and the transformed information of the principal components is inversely transformed to the 2 × M summary values again. Accordingly, based on the summary value of the sensor data acquired from the substrate processing apparatus 3, the information processing apparatus 1 can obtain a predicted value of the normal summary value for the sensor data.

[0070]

The information processing apparatus 1 performs, using predicted values of 2 × M summary values acquired from the prediction model 5, preprocessing on the sensor data including the original M × N values. In this example, the sensor data includes values related to M sensors 4 and N time series values for each of the sensors 4. The information processing apparatus 1 predicts the two summary values, that is, the average value and the standard deviation, for each sensor 4, and standardizes N time series values using predicted values of the average value and the standard deviation. The information processing apparatus 1 can perform standardization for the M sensors 4 in the same manner, and can accordingly standardize all of the M × N values in the sensor data.

[0071]

The prediction model 5 used in the preprocessing is generated by the information processing apparatus 1 before performing the machine learning for generating the learning model 2. The information processing apparatus 1 acquires the normal sensor data used for the machine learning of the learning model 2 and calculates the summary value for each sensor 4 with respect to the acquired sensor data. The information processing apparatus 1 generates the prediction model 5 that performs dimensionality reduction based on the calculated plurality of summary values by principal component analysis or the like, and restores the dimensionally reduced information to original dimensionality. Since the principal component analysis is a technique known in the related art, a detailed description thereof will be omitted. The prediction

model 5 is not limited to using the principal component analysis, and any learning model that can reduce dimensionality of input data and restore the data to original dimensionality may be adopted. The prediction model 5 may be any learning model as long as a normal summary value can be predicted for a received summary value.

[0072]

The information processing apparatus 1 performs the preprocessing using the prediction model 5 generated in advance both when the learning model 2 is generated by the machine learning and when the anomaly detection is performed using the trained learning model 2. The prediction model 5 used in any case is the same (having the same internal parameter determined by the machine learning). The information processing apparatus 1 generates the prediction model 5 in advance, and stores information such as the internal parameter of the prediction model 5 in the model information storage 12c.

[0073]

The information processing apparatus 1 according to the embodiment predicts, using the prediction model 5, the normal summary value from the summary value of the sensor data, and performs the preprocessing such as the standardization or the normalization using the predicted normal summary value. Accordingly, through preprocessing, the information processing apparatus 1 can be expected to reduce or eliminate a value change caused by a factor such as an individual difference or a change over time of the substrate processing apparatus 3, and thus improve the prediction accuracy of the learning model 2.

[0074]

<Summary>

In the information processing system according to the embodiment having the above configuration, the information processing apparatus 1 receives, as an input, the sensor data including the plurality of measurement values measured by the plurality of sensors 4 of the substrate processing apparatus 3, and generates, by the machine learning, the learning model 2 that outputs the sensor data including a plurality of corresponding normal measurement values. The information processing apparatus 1 acquires the normal sensor data related to the substrate processing from among the sensor data stored in the sensor data storage 12b, inputs the acquired sensor data into the learning model 2, and acquires the sensor data output by the learning model 2. The information processing apparatus 1 calculates the individual error for each measurement value in the sensor data input to the learning model 2 and the sensor data output by the learning model 2, and assigns weights to the calculated individual errors to calculate the total error. The information processing apparatus 1 performs the machine learning by updating the internal parameter of the learning model 2 based on the calculated total error, and thus generates the learning model 2. According to a convergence status of the individual errors, the information

processing apparatus 1 adjusts the weights assigned to the individual errors. Accordingly, the information processing system according to the embodiment can perform the machine learning in consideration of sensor data including a large amount of noise and sensor data including a small amount of noise to generate the learning model 2, for example, and an influence of noise or the like in the substrate processing state detection using the generated learning model 2 can be expected to be reduced.

[0075]

In the information processing system according to the embodiment, the information processing apparatus 1 calculates the weights based on formula (1) to formula (3) described above. The information processing apparatus 1 adjusts the weight to be assigned to each individual error such that, as an individual error at a certain time point becomes smaller than a previous individual error, a weight assigned to this individual error becomes larger. That is, the information processing apparatus 1 adjusts the weight to be assigned to each individual error such that a weight assigned to an individual error that is difficult to converge becomes smaller. Accordingly, the information processing system according to the embodiment can be expected to reduce machine learning importance for a sensor 4 where noise is large and whose individual error is difficult to converge, and thus generate the learning model 2.

[0076]

In the information processing system according to the embodiment, the learning model 2 generated by the information processing apparatus 1 has an autoencoder configuration. Accordingly, the information processing system according to the embodiment can perform the machine learning for generating the learning model 2 by machine learning using training data that does not have a ground truth label, that is, by unsupervised machine learning. It can be expected that the information processing apparatus 1 generates the learning model 2 for predicting the normal sensor data by performing the machine learning using the normal sensor data.

[0077]

In the information processing system according to the embodiment, the information processing apparatus 1 acquires the sensor data including a plurality of sensor values measured by the plurality of sensors 4 according to the substrate processing from the substrate processing apparatus 3. The information processing apparatus 1 inputs the acquired sensor data to the learning model 2 generated in advance by the machine learning described above, and acquires the sensor data output by the learning model 2. The information processing apparatus 1 detects a state such as an anomaly related to the substrate processing based on a comparison between the sensor data input to the learning model 2 and the sensor data output by the learning model 2. Accordingly, the information processing system according to the embodiment can be

expected to reduce influence of noise and the like in substrate processing state detection using the learning model 2.

[0078]

In the information processing system according to the embodiment, the information processing apparatus 1 stores, in the model information storage 12c, information on the weight for each sensor data measurement value calculated at the time of the machine learning for generating the learning model 2. At the time of state detection, the information processing apparatus 1 determines, based on the stored weight, a sensor to be excluded from the state detection. Accordingly, the information processing system according to the embodiment can exclude, from the state detection, a sensor 4 whose individual error is difficult to converge at the time of the machine learning and whose learning importance is thus reduced, and thus can be expected to prevent erroneous detection of a state of the sensor 4 having low prediction accuracy by the learning model 2.

[0079]

The embodiments disclosed herein are exemplary in all respects and can be considered to be not restrictive. The scope of the present disclosure is indicated by the claims, not the above-described meaning, and is intended to include all modifications within the meaning and scope equivalent to the claims.

[0080]

The features described in each embodiment can be combined with each other. In addition, the independent and dependent claims set forth in the claims can be combined with each other in any and all combinations, regardless of the reciting format. Furthermore, the claims use a format of describing claims that recite two or more other claims (multi-claim format). However, the present disclosure is not limited thereto. The claims may also be described using a format of multi-claims reciting at least one multi-claim (multi-multi claims).

Reference Signs List

[0081]

1 information processing apparatus (computer)

2 learning model

3 substrate processing apparatus

4 sensor

5 prediction model

11 processor

11a sensor data acquisition unit

11b learning processor

11c anomaly detector

11d display processor

12 storage

12a program (computer program)

12b sensor data storage

12c model information storage

13 communication unit

14 display unit

15 operation unit

CLAIMS

[Claim 1]

A computer program for causing a computer to execute processing of generating, by machine learning, a learning model that receives, as an input, sensor data including a plurality of measurement values measured by a plurality of sensors related to substrate processing, and outputs sensor data including a plurality of corresponding measurement values, the computer program causing the computer to execute processing of

acquiring the sensor data related to the substrate processing,

inputting the acquired sensor data into the learning model and acquiring the sensor data output by the learning model,

calculating an individual error for each measurement value between the sensor data input into the learning model and the sensor data output by the learning model,

assigning weights to a plurality of the calculated individual errors, respectively, to calculate a total error,

updating an internal parameter of the learning model based on the calculated total error, and

adjusting the weights assigned to the individual errors according to a convergence status of the individual errors.

[Claim 2]

The computer program according to claim 1, wherein

the weights assigned to the individual errors are adjusted such that, as an individual error at one time point becomes smaller than an individual error at a time point before the one time point, a weight assigned to the individual error at the one time point becomes larger.

[Claim 3]

The computer program according to claim 1, wherein

the learning model is an autoencoder.

[Claim 4]

The computer program according to claim 1, wherein

based on the acquired sensor data, the measurement values in the sensor data are changed within a predetermined range to generate augmented sensor data, and

the machine learning of the learning model is performed using the generated augmented sensor data as the sensor data.

[Claim 5]

The computer program according to claim 1, wherein

a summary value of the acquired sensor data is calculated,

a prediction model for predicting a summary value of the sensor data based on the calculated summary value is generated,

preprocessing is performed on the sensor data using the summary value predicted by the generated prediction model, and

the sensor data subjected to the preprocessing is input into the learning model.

[Claim 6]

A computer program causing a computer to execute processing of

acquiring sensor data including a plurality of measurement values measured by a plurality of sensors related to substrate processing,

inputting, into a learning model that is subjected to machine learning to receive the sensor data as an input and output corresponding sensor data, the acquired sensor data, and acquiring the sensor data output by the learning model, and

detecting a state related to the substrate processing based on a comparison between the sensor data input into the learning model and the sensor data output by the learning model, wherein

the machine learning of the learning model is performed by

inputting the sensor data into the learning model and acquiring the sensor data output by the learning model,

calculating an individual error for each measurement value between the sensor data input into the learning model and the sensor data output by the learning model,

assigning weights to a plurality of the calculated individual errors, respectively, to calculate a total error,

updating an internal parameter of the learning model based on the calculated total error, and

adjusting the weights assigned to the individual errors according to a convergence status of the individual errors.

[Claim 7]

The computer program according to claim 6, wherein

information on a weight for each measurement value in the sensor data when the machine learning is performed is acquired, and

a sensor to be excluded from state detection is determined based on the acquired weight.

[Claim 8]

The computer program according to claim 7, wherein

an anomaly degree of each sensor is calculated based on the comparison between the sensor data input into the learning model and the sensor data output by the learning model, and

information on one or a plurality of sensors targeted for the state detection is output in association with the anomaly degree of each of the sensors.

[Claim 9]

The computer program according to claim 8, wherein

information on a plurality of sensors is output in an order corresponding to the anomaly degree.

[Claim 10]

The computer program according to claim 7, wherein

information on a sensor targeted for the state detection and information on the sensor to be excluded from the state detection are output in a distinguishable manner.

[Claim 11]

The computer program according to claim 6, wherein

a summary value of the acquired sensor data is calculated,

based on the calculated summary value, a summary value of the sensor data is predicted using a prediction model subjected to machine learning to receive the summary value of the sensor data as an input and predict a summary value of the sensor data,

preprocessing for the acquired sensor data is performed using the predicted summary value, and

the sensor data subjected to the preprocessing is input into the learning model.

[Claim 12]

A learning model generation method for an information processing apparatus to generate, by machine learning, a learning model that receives, as an input, sensor data including a plurality of measurement values measured by a plurality of sensors related to substrate processing, and outputs sensor data including a plurality of corresponding measurement values, the learning model generation method comprising:

acquiring the sensor data related to the substrate processing;

inputting the acquired sensor data into the learning model and acquiring the sensor data output by the learning model;

calculating an individual error for each measurement value between the sensor data input into the learning model and the sensor data output by the learning model;

assigning weights to a plurality of the calculated individual errors, respectively, to calculate a total error;

updating an internal parameter of the learning model based on the calculated total error; and

adjusting the weights assigned to the individual errors according to a convergence status of the individual errors.

[Claim 13]

An information processing method comprising:

causing a computer to execute, by an information processing apparatus, processing of

acquiring sensor data including a plurality of measurement values measured by a plurality of sensors related to substrate processing,

inputting, into a learning model that is subjected to machine learning to receive the sensor data as an input and output corresponding sensor data, the acquired sensor data, and acquiring the sensor data output by the learning model, and

detecting a state related to the substrate processing based on a comparison between the sensor data input into the learning model and the sensor data output by the learning model, wherein

the machine learning of the learning model is performed by

inputting the sensor data into the learning model and acquiring the sensor data output by the learning model,

calculating an individual error for each measurement value between the sensor data input into the learning model and the sensor data output by the learning model,

assigning weights to a plurality of the calculated individual errors, respectively, to calculate a total error,

updating an internal parameter of the learning model based on the calculated total error, and

adjusting the weights assigned to the individual errors according to a convergence status of the individual errors.

[Claim 14]

An information processing apparatus comprising:

a processor configured to generate, by machine learning, a learning model that receives, as an input, sensor data including a plurality of measurement values measured by a plurality of sensors related to substrate processing, and outputs sensor data including a plurality of corresponding measurement values, wherein

the processor

acquires the sensor data related to the substrate processing,

inputs the acquired sensor data into the learning model and acquires the sensor data output by the learning model,

calculates an individual error for each measurement value between the sensor data input into the learning model and the sensor data output by the learning model,

assigns weights to a plurality of the calculated individual errors, respectively, to calculate a total error,

updates an internal parameter of the learning model based on the calculated total error, and

adjusts the weights assigned to the individual errors according to a convergence status of the individual errors.

[Claim 15]

An information processing apparatus comprising:

a processor, wherein

the processor causes a computer to execute processing of

acquiring sensor data including a plurality of measurement values measured by a plurality of sensors related to substrate processing,

inputting, into a learning model that is subjected to machine learning to receive sensor data as an input and output corresponding sensor data, the acquired sensor data, and acquiring the sensor data output by the learning model, and

detecting a state related to the substrate processing based on a comparison between the sensor data input into the learning model and the sensor data output by the learning model, and

the machine learning of the learning model is performed by

inputting the sensor data into the learning model and acquiring the sensor data output by the learning model,

calculating an individual error for each measurement value between the sensor data input into the learning model and the sensor data output by the learning model,

assigning weights to a plurality of the calculated individual errors,

respectively, to calculate a total error,

updating an internal parameter of the learning model based on the calculated total error, and

adjusting the weights assigned to the individual errors according to a convergence status of the individual errors.