(12) 特許協力条約に基づいて公開された国際出願

(19) 世界知的所有権機関
国際事務局

(43) 国際公開日
2020年9月17日(17.09.2020)

WIPO | PCT

(10) 国際公開番号
WO 2020/184561 A1

(51) 国際特許分類:
G06N 20/00 (2019.01)

(21) 国際出願番号： PCT/JP2020/010304

(22) 国際出願日： 2020年3月10日(10.03.2020)

(25) 国際出願の言語： 日本語

(26) 国際公開の言語： 日本語

(30) 優先権データ：
特願 2019-045556 2019年3月13日(13.03.2019) JP

(71) 出願人: 日本電信電話株式会社 (**NIPPON TELEGRAPH AND TELEPHONE CORPORATION**) [JP/JP]; 〒1008116 東京都千代田区大手町一丁目５番１号 Tokyo (JP).

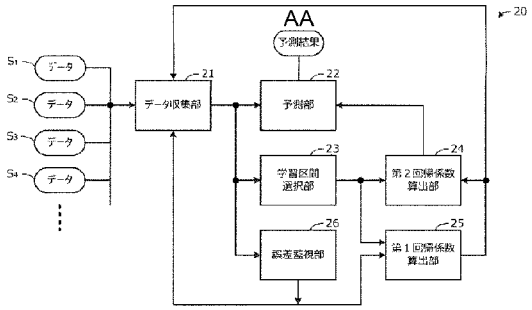
(72) 発明者: 高嶋 洋一 (**TAKASHIMA, Yoichi**); 〒1808585 東京都武蔵野市緑町三丁目９番１１号 ＮＴＴ 知的財産センタ内 Tokyo (JP). 湯口 昌宏(**YUGUCHI, Masahiro**); 〒1808585 東京都武蔵野市緑町三丁目９番１１号 ＮＴＴ 知的財産センタ内 Tokyo (JP). 山田 智広(**YAMADA, Tomohiro**); 〒1808585 東京都武蔵野市緑町三丁目９番１１号 ＮＴＴ 知的財産センタ内 Tokyo (JP).

(74) 代理人: 特許業務法人太陽国際特許事務所 (**TAIYO, NAKAJIMA & KATO**); 〒1600022 東京都新宿区新宿４丁目３番１７号 Tokyo (JP).

(81) 指定国(表示のない限り、全ての種類の国内保護が可能): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL,

(54) **Title:** DATA PREDICTION DEVICE, DATA PREDICTION METHOD, AND DATA PREDICTION PROGRAM

(54) 発明の名称： データ予測装置、データ予測方法、及びデータ予測プログラム



21 Data collection unit
22 Prediction unit
23 Learning segment selection unit
24 Second regression coefficient computation unit
25 First regression coefficient computation unit
26 Error monitoring unit
S1, S2, S3, S4 Data
AA Prediction result

(57) **Abstract:** The purpose of the present invention is to reduce the quantity of data used to calculate regression coefficients, reduce the bandwidth for collecting and transmitting data, and reduce computation throughput in a learning process. A data prediction device according to the present disclosure includes: a data collection unit that transmits a collected plurality of data, and selected data that was selected from the plurality of data; a first regression coefficient computation unit that computes regression coefficients on the basis of the received plurality of data and using a first computation method for calculating regression coefficients by L1 regularization, and transmits selection information for selecting data corresponding to regression coefficients among the calculated regression coefficients having an absolute value equal to or greater than a threshold value to the data collection unit; a second regression coefficient computation unit that computes regression coefficients on the basis of the received selected data and using a second computation method different from the first computation method; and a prediction unit that outputs a prediction result that was predicted on the basis of the received selected data and the regression coefficients computed by the second regression coefficient computation unit.

WO 2020/184561 A1

[続葉有]

ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW.

(84) 指定国(表示のない限り、全ての種類の広域保護が可能): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), ユーラシア (AM, AZ, BY, KG, KZ, RU, TJ, TM), ヨーロッパ (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

添付公開書類:

一 国際調査報告(条約第21条(3))

(57) 要約:回帰係数を求める際に使用するデータの数を低減し、データを収集及び伝送する回線容量を低減し、学習過程における計算処理量を低減する、ことを目的とする。 本開示のデータ予測装置は、収集した複数のデータ、及び複数のデータから選択した選択データを送信するデータ収集部と、Ｌ１正則化により回帰係数を求める第１算出法を用いて、受信した前記複数のデータに基づいて回帰係数を算出し、算出した回帰係数のうち絶対値が閾値以上の回帰係数に対応するデータを選択するための選択情報を前記データ収集部に送信する第１回帰係数算出部と、第１算出法と異なる第２算出法を用いて、受信した前記選択データに基づいて回帰係数を算出する第２回帰係数算出部と、受信した前記選択データ、及び前記第２回帰係数算出部で算出された回帰係数に基づいて予測した予測結果を出力する予測部と、を含む。

明　細　書

発明の名称 ：

データ予測装置、データ予測方法、及びデータ予測プログラム

技術分野

[0001] 本開示は、データ予測装置、データ予測方法、及びデータ予測プログラムに関する。

背景技術

[0002] 各所に設置した大量のセンサから取得されるデータから、その中の特定のデータの近い将来の値を予測する問題では、（１）式に例示するように、図１に示す線形回帰による予測が行われることがある。下記（１）式は、図１に例示する１～４の４つのセンサの時刻ｔ－３から時刻ｔまでの単位時間毎のデータのうち時刻ｔ－２から時刻ｔまでの１２個のデータからセンサ１の時刻ｔ＋１におけるデータ$D_{1,t+1}$を予測する。

$$\beta_0+\beta_1 D_{1,t-2}+\beta_2 D_{2,t-2}+\beta_3 D_{3,t-2}+\beta_4 D_{4,t-2}+$$
$$\beta_5 D_{1,t-1}+\beta_6 D_{2,t-1}+\beta_7 D_{3,t-1}+\beta_8 D_{4,t-1}+$$
$$\beta_9 D_{1,t}+\beta_{10} D_{2,t}+\beta_{11} D_{3,t}+\beta_{12} D_{4,t}=D_{1,t+1} \quad \cdots (1)$$

[0003] β_k（k＝０，…，１２）は回帰係数であり、$D_{l,t-m}$は、図１に円で例示するデータであり、lはセンサの番号、ｔ－ｍは時刻を表し、ｍは自然数を表す。

[0004] 詳細には、以下の学習過程及び予測過程によりデータの予測が行われる。まず、予測するデータに関係しそうな各センサの過去のデータを収集して線形回帰式を作成し、予測しようとする現象（予測対象の特異な動き）が生じているサンプル区間のデータを線形回帰式の選択変数に代入して、回帰係数β_kを連立方程式により求める(学習過程)。次に、求められた回帰係数β_kにより、センサのデータから連続的に予測値を計算する(予測過程)。

[0005] 非常に多量のセンサからの過去の長い期間のデータを用いる場合には、サンプル区間のデータが不足し、連立方程式の解が求まらなくなる場合がある

。例えば、河川の水位の予想を行う場合、センサからのデータの代わり、あるいは、センサからのデータに追加して、例えば、気象庁で作成されるメッシュ状の雨量の予報値を用いることができる。しかしながら、この場合はさらに回帰係数の数が増えることになり、サンプル区間のデータが不足する場合がある。

[0006] サンプル区間のデータが不足した場合、Ｌ１正則化という方法（非特許文献１）を用いて、（２）式に例示するように、近似解の回帰係数を求めることができる。このＬ１正則化には、予測に貢献しにくい回帰係数をゼロに近づける性質がある。

[数1]

$$\hat{\beta} = \arg\min \sum_{i=1}^{N} \left(y_i - \sum_j \beta_j x_{ij} \right)^2 \quad \cdots (2)$$

$$\sum_j |\beta_j| \leq t_c$$

y_iは目的変数であり、x_{ij}は説明変数であり、t_cは調整パラメータである。

[0007] Ｌ１正則化を用いて構成される、関連技術のデータ予測装置１０を図２に示す。データ収集部１１は多数のセンサの各々からデータを収集し、予測部１２及び学習区間選択部１３に同じデータを送信する。学習区間選択部１３では、予測しようとする事象が現れていそうなデータ区間内のデータをルールベースで選択し、選択したデータを学習するべきデータとして回帰係数算出部１４に送信する。回帰係数算出部１４では、Ｌ１正則化を用いて回帰係数を求め、当該回帰係数を予測部１２に送信する。予測部１２は、回帰係数算出部１４から受信した回帰係数と、データ収集部１１から受信したデータ

と、を線形回帰式に代入して連続的に予測値を算出する。

先行技術文献

非特許文献

[0008] 非特許文献1：Robert Tibshirani, "Regression Shrinkage and Selection via the Lasso", Journal of the Royal Statistical Society, Series B (Methodological), 1996, Vol. 58, No. 1, pp.267-288

発明の概要

発明が解決しようとする課題

[0009] Ｌ１正則化により、多くの説明変数を使用する場合にも回帰係数を求めることはできるようになるが、予測に影響を及ぼすセンサの範囲、及びデータ計測時間の範囲は事前に既知でないため、広めに範囲を設定することになる。そのため、入力として使用するデータの数が非常に多くなり、データを収集及び伝送する回線容量を圧迫し、また、学習過程における計算処理量も膨大になる、という問題が発生する。

[0010] 本開示では、回帰係数を求める際に使用するデータの数を低減し、データを収集及び伝送する回線容量を低減し、学習過程における計算処理量を低減する、ことを目的とする。

課題を解決するための手段

[0011] 本開示の第１態様のデータ予測装置は、収集した複数のデータ、及び受信した選択情報に基づいて前記複数のデータから選択した選択データを送信するデータ収集部と、Ｌ１正則化により回帰係数を求める第１算出法を用いて、前記データ収集部から受信した前記複数のデータに基づいて回帰係数を算出し、算出した回帰係数のうち絶対値が閾値以上の回帰係数に対応するデータを選択するための選択情報を前記データ収集部に送信する第１回帰係数算出部と、前記第１算出法と異なる第２算出法を用いて、前記データ収集部から受信した前記選択データに基づいて回帰係数を算出する第２回帰係数算出部と、前記データ収集部から受信した前記選択データ、及び前記第２回帰係

数算出部で算出された回帰係数に基づいて予測した予測結果を出力する予測部と、を含む。

[0012] 本開示の第２態様は、第１態様のデータ予測装置であり、前記予測結果と実測値との誤差であって、所定値以上の前記誤差の発生回数が所定回数以上になった場合に、前記第２回帰係数算出部に対して前記第２算出法による回帰係数の再算出の指示を行う誤差監視部をさらに含む。

[0013] 本開示の第３態様は、第２態様のデータ予測装置であって、前記誤差監視部は、前記再算出の指示を行った後に、前記誤差の発生回数が再度前記所定回数以上になった場合に、前記第１回帰係数算出部に対して、前記第１算出法による回帰係数の再算出の指示、及び前記データ収集部に対する選択情報の再送信の指示を行う。

[0014] 本開示の第４態様のデータ予測装置は、収集した複数のデータ、及び受信した選択情報に基づいて前記複数のデータから選択した選択データを送信するデータ収集部と、Ｌ１正則化により回帰係数を求める算出法を用いて、前記データ収集部から受信した前記複数のデータに基づいて回帰係数を算出し、算出した回帰係数のうち非ゼロの回帰係数に対応するデータを選択するための選択情報を前記データ収集部に送信する回帰係数算出部と、前記データ収集部から受信した前記選択データ、及び前記回帰係数算出部で算出された前記非ゼロの回帰係数に基づいて予測した予測結果を出力する予測部と、を含む。

[0015] 本開示の第５態様は、第４態様のデータ予測装置であって、前記予測結果と実測値との誤差であって、所定値以上の前記誤差の発生回数が所定回数以上になった場合に、前記回帰係数算出部に対して前記算出法による回帰係数の再算出の指示を行う誤差監視部をさらに含む。

[0016] 本開示の第６態様のプログラムは、収集した複数のデータ、及び受信した選択情報に基づいて前記複数のデータから選択した選択データを送信し、Ｌ１正則化により回帰係数を求める第１算出法を用いて、受信した前記複数のデータに基づいて回帰係数を算出し、算出した回帰係数のうち絶対値が閾値

以上の回帰係数に対応するデータを選択するための選択情報を送信し、前記第１算出法と異なる第２算出法を用いて、受信した前記選択データを使用して回帰係数を算出し、受信した前記選択データ、及び前記第２算出法を用いて算出された回帰係数に基づいて予測した予測結果を出力する、データ予測処理をコンピュータに実行させる。

[0017] 本開示の第７態様のデータ予測方法では、コンピュータが、収集した複数のデータ、及び受信した選択情報に基づいて前記複数のデータから選択した選択データを送信し、Ｌ１正則化により回帰係数を求める第１算出法を用いて、受信した前記複数のデータに基づいて回帰係数を算出し、算出した回帰係数のうち絶対値が閾値以上の回帰係数に対応するデータを選択するための選択情報を送信し、前記第１算出法と異なる第２算出法を用いて、受信した前記選択データを使用して回帰係数を算出し、受信した前記選択データ、及び前記第２算出法を用いて算出された回帰係数に基づいて予測した予測結果を出力する。

発明の効果

[0018] 本開示では、データの数を低減し、データを収集及び伝送する回線容量を低減し、学習過程における計算処理量を低減する、ことができる。

図面の簡単な説明

[0019] [図1]線形回帰式によるデータの予測を説明する模式図である。

[図2]関連技術のデータ予測装置を例示するブロック図である。

[図3]本実施形態のデータ予測装置を例示するブロック図である。

[図4]本実施形態のデータ予測装置のハードウェア構成図を例示するブロック図である。

[図5]本実施形態の学習フェーズの処理の流れを例示する模式図である。

[図6]本実施形態の予測フェーズの処理の流れを例示する模式図である。

[図7]本実施形態の再選択フェーズの処理の流れを例示する模式図である。

[図8]本実施形態のデータ予測装置を例示するブロック図である。

[図9]本実施形態のデータ予測装置の変形例を例示するブロック図である。

発明を実施するための形態

[0020] 本実施形態のデータ予測装置２０を図３に例示する。データ収集部２１は多数のセンサS_1，S_2，S_3，…の各々からデータを収集し、予測部２２及び学習区間選択部２３に同じデータを送信する。データ収集部２１が、予測部２２及び学習区間選択部２３に送信するデータとしては、多数のセンサS_1，S_2，S_3，…から収集した複数のデータ、及び後述する第１回帰係数算出部２５から受信した選択情報に基づいて収集した複数のデータから選択した選択データがある。

[0021] 学習区間選択部２３では、予測しようとする事象が現れていることが予測されるデータ区間をルールベースで選択し、学習するべきデータを第１回帰係数算出部２５及び第２回帰係数算出部２４に送信する。すなわち、学習区間選択部２３は、予測しようとする事象が現れていないことが予測される区間に存在する不要なデータ及び変化の乏しいデータを取り除く。

[0022] 第１回帰係数算出部２５は、Ｌ１正則化により近似解の回帰係数を求める第１算出法を用いて、データ収集部２１から送信され、かつ学習区間選択部２３で選択された複数のデータに基づいて回帰係数を算出し、算出した回帰係数のうち絶対値が閾値以上の回帰係数を選択し、選択した回帰係数を示す情報を、データ収集部２１及び第２回帰係数算出部２４の各々に送信する。

[0023] データ収集部２１では、第１回帰係数算出部２５から受信した情報に基づいて、収集した複数のデータから絶対値が閾値以上の回帰係数に対応するデータを選択し、選択したデータを選択データとして予測部２２、学習区間選択部２３、及び誤差監視部２６の各々に送信する。第１回帰係数算出部２５からデータ収集部２１に送信される情報は、データ収集部２１においてデータを選択するために使用されるので、以下では選択情報という。

[0024] なお、データ収集部２１では、絶対値が閾値以上の回帰係数に対応するデータが選択データとして選択され、絶対値が閾値未満の回帰係数に対応するデータが削除される。このため、データ収集部２１に送信する選択情報としては、絶対値が閾値以上の回帰係数を示す情報に代えて、絶対値が閾値未満

の回帰係数を示す情報を用い、データ収集部２１において、絶対値が閾値未満の回帰係数に対応するデータを削除することにより、選択データを得るようにしてもよい。

[0025] 第２回帰係数算出部２４は、データ収集部２１から送信され、かつ学習区間選択部２３を通過した選択データを使用して、Ｌ１正則化ではなく、Ｌ２正則化または重回帰を用いて回帰係数を求める。

[0026] 第２回帰係数算出部２４で算出された回帰係数は、Ｌ２正則化または重回帰を用いた回帰計算による予測が行われる予測部２２に送信される。データ収集部２１では、選択データを予測部２２及び学習区間選択部２３に送信する。すなわち、データ収集部２１は、全てのセンサS_1，S_2，S_3，…の各々からデータを収集するが、データの送信は、選択的に行う。

[0027] 予測部２２では、データ収集部２１から受信した選択データと、第２回帰係数算出部２４から受信した回帰係数とを用いて、線形回帰式で予測した予測結果を出力する。

[0028] すなわち、予測結果は、絶対値が閾値未満の回帰係数に対応するデータが削除され、絶対値が閾値以上の回帰係数と絶対値が閾値以上の回帰係数に対応するデータとから算出されるため、学習過程で使用するデータ量が低減され、計算処理量を低減することができる。

[0029] 誤差監視部２６では、予測部２２による予測結果と、データ収集部２１で収集されたデータである実測値との差である誤差を監視し、所定値以上の誤差が所定回数(許容回数)以上発生した場合に、データ収集部２１と第１回帰係数算出部２５に、再選択指示を送信する。データ収集部２１は、この再選択指示により、選択データの再選択が終わるまでの一定期間、選択データではなく、全データを後段に送信するようになり、第１回帰係数算出部２５では、再度、全データを対象にしたＬ１正則化によるデータの選択が行われる。この際、Ｌ２正則化または重回帰による予測結果の誤差が、Ｌ１正則化による予測結果の誤差よりも大きい場合にのみ、選択データの再選択を行い、Ｌ２正則化または重回帰による予測結果の誤差が、Ｌ１正則化による予測結

果の誤差以下である場合には、再選択を行わない。

[0030] なお、誤差監視部２６では、予測部２２による予測結果と、実測値の差である誤差を監視し、所定値以上の誤差が所定回数(許容回数)以上発生した場合に、第２回帰係数算出部２４で、回帰係数の再算出を行うようにしてもよい。回帰係数の再算出後、所定値以上の誤差が所定回数（許容回数）以上発生した場合に、データ収集部２１と第１回帰係数算出部２５に、再選択指示を送信するようにしてもよい。

[0031] 図４に、データ予測装置２０のハードウェア構成を例示する。データ予測装置２０は、一例として、図４に示すように、ＣＰＵ（Central Processing Unit）５１、一次記憶部５２、二次記憶部５３、及び、外部インタフェース５４を含む。ＣＰＵ５１は、ハードウェアであるプロセッサの一例である。ＣＰＵ５１、一次記憶部５２、二次記憶部５３、及び、外部インタフェース５４は、バス５９を介して相互に接続されている。

[0032] 一次記憶部５２は、例えば、ＲＡＭ（Random Access Memory）などの揮発性のメモリである。二次記憶部５３は、例えば、ＨＤＤ（Hard Disk Drive）、又はＳＳＤ（Solid State Drive）などの不揮発性のメモリである。

[0033] 二次記憶部５３は、プログラム格納領域５３Ａ及びデータ格納領域５３Ｂを含む。プログラム格納領域５３Ａは、一例として、データ予測プログラムなどのプログラムを記憶している。データ格納領域５３Ｂは、一例として、センサからのデータ及びデータ予測処理中の中間データなどを記憶する。

[0034] ＣＰＵ５１は、プログラム格納領域５３Ａからデータ予測プログラムを読み出して一次記憶部５２に展開する。ＣＰＵ５１は、データ予測プログラムをロードして実行することで、図３のデータ収集部２１、予測部２２、学習区間選択部２３、第２回帰係数算出部２４、第１回帰係数算出部２５、及び誤差監視部２６として動作する。

[0035] なお、データ予測プログラムなどのプログラムは、外部サーバに記憶され、ネットワークを介して、一次記憶部６２に展開されてもよい。また、データ予測プログラムなどのプログラムは、Digital Versatile Disc（ＤＶＤ）

などの非一時的記録媒体に記憶され、記録媒体読込装置を介して、一次記憶部５２に展開されてもよい。

[0036] 外部インタフェース５４には外部装置が接続され、外部インタフェース５４は、外部装置とＣＰＵ５１との間の各種情報の送受信を司る。図４では、外部インタフェース５４に、センサ３１Ａ、及び危険報知システム３１Ｂが接続されている例を示している。なお、センサ３１Ａは、多数のセンサを含む。

[0037] データ予測装置２０で予測されたデータは、例えば、外部インタフェース５４に接続される危険報知システム３１Ｂに送信され、危険報知システム３１Ｂにおける危険報知処理に使用されてもよい。また、データ予測装置２０で予測されたデータは、例えば、外部インタフェース５４に接続される外部記憶装置に記録されてもよいし、外部インタフェース５４に接続されるディスプレイの画面に文字または画像として表示されてもよい。

[0038] また、データ予測装置２０は、専用装置であってもよいし、ワークステーション、パーソナルコンピュータ、またはタブレットなどの汎用装置であってもよい。

[0039] 図５～図７に、本実施形態のデータ予測処理を例示する。図５は、学習フェーズの流れを例示する。

[0040] データ収集部２１は、手順Ｓ１０１で、例えば、多数のセンサを含むセンサ３１Ａからデータを収集する。データ収集部２１は、手順Ｓ１０２で、学習区間選択部２３に、収集したデータを送信する。

[0041] 学習区間選択部２３は、手順Ｓ１０３で、予測に適切そうなデータ区間内のデータを選択して第２回帰係数算出部２４と第１回帰係数算出部２５に送信する。第１回帰係数算出部２５は、Ｌ１正則化を用いて回帰係数を算出し、絶対値が閾値以上の回帰係数を選択して、第２回帰係数算出部２４とデータ収集部２１に送信する。

[0042] 第２回帰係数算出部２４は、第１回帰係数算出部２５で選択されたデータによる通常の回帰計算、例えば、Ｌ２正則化または重回帰により、回帰係数

を求めて予測部２２に送信する。

[0043] 図６は、予測フェーズの流れを例示する。データ収集部２１は、手順Ｓ１０６で、手順Ｓ１０４で選択されたデータを予測部２２及び誤差監視部２６に送信する。予測部２２は、手順Ｓ１０７で、手順Ｓ１０５で求めた回帰係数とデータとに基づいて予測結果を算出し、誤差監視部２６に送信すると共に、例えば、外部記憶装置に記憶する。

[0044] 図７は、誤差監視及び再選択フェーズの流れを例示する。誤差監視部２６は、手順Ｓ１０８で、予測結果とデータ（実測値）との予測誤差（誤差）を算出し、所定値以上の予測誤差が所定回数（許容回数）以上生じた場合、第１回帰係数算出部２５に再選択指示を送信する。また、誤差監視部２６は、手順Ｓ１０８で、再選択指示を送信した場合、データ収集部２１に、再選択が完了するまで、センサ等から収集されたデータを全て後段（学習区間選択部２３、誤差監視部２６及び予測部２２）に送信するよう指示する。

[0045] 本実施形態によれば、Ｌ１正則化を予測に用いる場合より、予測誤差が大きくなる傾向を示す場合があるが、予測誤差が大きい場合、あるいは、定期的に、データの再選択を行うことにより、予測誤差を所定範囲内に収めることが可能である。

[0046] なお、データ予測装置２０は、図８に例示するように、データ収集装置３２Ａ、学習装置３２Ｂ及び予測装置３２Ｃを含んでいてもよい。データ収集装置３２Ａは、データ収集部２１を含み、学習装置３２Ｂは、学習区間選択部２３、第２回帰係数算出部２４、第１回帰係数算出部２５及び誤差監視部２６を含み、予測装置３２Ｃは予測部２２を含む。

[0047] センサ３１Ａとデータ収集部２１との間、及び予測部２２と、例えば、危険報知システム３１Ｂなどの予測結果の出力先との間は、伝送回線で接続されている。また、図８に示す例では、データ収集部２１と学習区間選択部２３との間、第２回帰係数算出部２４と予測部２２との間も伝送回線で接続されている。

[0048] 図９に、本実施形態の変形例のデータ予測装置４０を例示する。データ予

測装置２０と同様の構成及び作用については、説明を適宜省略する。データ予測装置４０は、第２回帰係数算出部２４及び第１回帰係数算出部２５に代えて、回帰係数算出部４４を含む点で、図３のデータ予測装置２０と異なる。データ予測装置４０の回帰係数算出部４４は、Ｌ１正則化を使用して回帰係数を算出する。回帰係数算出部４４は、非ゼロの回帰係数をデータ収集部４１及び予測部４２に送信する。

[0049] データ収集部４１では、回帰係数算出部４４から受信した情報に基づいて、収集した複数のデータから非ゼロの回帰係数に対応するデータを選択し、選択したデータを選択データとして予測部４２、学習区間選択部４３、及び誤差監視部４６の各々に送信する。

[0050] データ予測装置４０では、非ゼロの回帰係数に対応するデータを選択し、ゼロの回帰係数に対応するデータを選択しないことによって、データの選択が、回帰係数の算出と共に行われていると見なしている。

[0051] 変形例では、データの選択と回帰係数の算出を同時に行うことが可能となるため、計算コストを低減することができる。

[0052] 本開示のデータ予測装置は、収集した複数のデータ、及び受信した選択情報に基づいて前記複数のデータから選択した選択データを送信するデータ収集部と、Ｌ１正則化により回帰係数を求める第１算出法を用いて、前記データ収集部から受信した前記複数のデータに基づいて回帰係数を算出し、算出した回帰係数のうち絶対値が閾値以上の回帰係数に対応するデータを選択するための選択情報を前記データ収集部に送信する第１回帰係数算出部と、前記第１算出法と異なる第２算出法を用いて、前記データ収集部から受信した前記選択データに基づいて回帰係数を算出する第２回帰係数算出部と、前記データ収集部から受信した前記選択データ、及び前記第２回帰係数算出部で算出された回帰係数に基づいて予測した予測結果を出力する予測部と、を含む。

[0053] 本開示では、これにより、データの数を低減し、データを収集及び伝送する回線容量を低減し、学習過程における計算処理量を低減する、ことができ

る。なお、データ収集部２１から後段に送信するデータの数を低減する代わりに、データ収集部２１がデータを受信するセンサの数を低減してもよい。

[0054] 本開示では、第２回帰係数算出部２５での計算が、通常の回帰計算であるため、計算コストを低減することができ、計算リソースの少ない環境における処理を可能にする。

符号の説明

[0055] ２１ データ収集部

２２ 予測部

２６ 誤差監視部

２４ 第２回帰係数算出部

２５ 第１回帰係数算出部

３１Ａ センサ

５１ ＣＰＵ

５２ 一次記憶部

５３ 二次記憶部

請求の範囲

[請求項1] 収集した複数のデータ、及び受信した選択情報に基づいて前記複数のデータから選択した選択データを送信するデータ収集部と、

Ｌ１正則化により回帰係数を求める第１算出法を用いて、前記データ収集部から受信した前記複数のデータに基づいて回帰係数を算出し、算出した回帰係数のうち絶対値が閾値以上の回帰係数に対応するデータを選択するための選択情報を前記データ収集部に送信する第１回帰係数算出部と、

前記第１算出法と異なる第２算出法を用いて、前記データ収集部から受信した前記選択データに基づいて回帰係数を算出する第２回帰係数算出部と、

前記データ収集部から受信した前記選択データ、及び前記第２回帰係数算出部で算出された回帰係数に基づいて予測した予測結果を出力する予測部と、

を含むデータ予測装置。

[請求項2] 前記予測結果と実測値との誤差であって、所定値以上の前記誤差の発生回数が所定回数以上になった場合に、前記第２回帰係数算出部に対して前記第２算出法による回帰係数の再算出の指示を行う誤差監視部をさらに含む、

請求項１に記載のデータ予測装置。

[請求項3] 前記誤差監視部は、前記再算出の指示を行った後に、前記誤差の発生回数が再度前記所定回数以上になった場合に、前記第１回帰係数算出部に対して、前記第１算出法による回帰係数の再算出の指示、及び前記データ収集部に対する選択情報の再送信の指示を行う、

請求項２に記載のデータ予測装置。

[請求項4] 収集した複数のデータ、及び受信した選択情報に基づいて前記複数のデータから選択した選択データを送信するデータ収集部と、

Ｌ１正則化により回帰係数を求める算出法を用いて、前記データ収

集部から受信した前記複数のデータに基づいて回帰係数を算出し、算出した回帰係数のうち非ゼロの回帰係数に対応するデータを選択するための選択情報を前記データ収集部に送信する回帰係数算出部と、

前記データ収集部から受信した前記選択データ、及び前記回帰係数算出部で算出された前記非ゼロの回帰係数に基づいて予測した予測結果を出力する予測部と、

を含むデータ予測装置。

[請求項5] 前記予測結果と実測値との誤差であって、所定値以上の前記誤差の発生回数が所定回数以上になった場合に、前記回帰係数算出部に対して前記算出法による回帰係数の再算出の指示を行う誤差監視部をさらに含む、

請求項４に記載のデータ予測装置。

[請求項6] 収集した複数のデータ、及び受信した選択情報に基づいて前記複数のデータから選択した選択データを送信し、

Ｌ１正則化により回帰係数を求める第１算出法を用いて、受信した前記複数のデータに基づいて回帰係数を算出し、算出した回帰係数のうち絶対値が閾値以上の回帰係数に対応するデータを選択するための選択情報を送信し、

前記第１算出法と異なる第２算出法を用いて、受信した前記選択データを使用して回帰係数を算出し、

受信した前記選択データ、及び前記第２算出法を用いて算出された回帰係数に基づいて予測した予測結果を出力する、

データ予測処理をコンピュータに実行させるためのプログラム。

[請求項7] コンピュータが、

収集した複数のデータ、及び受信した選択情報に基づいて前記複数のデータから選択した選択データを送信し、

Ｌ１正則化により回帰係数を求める第１算出法を用いて、受信した前記複数のデータに基づいて回帰係数を算出し、算出した回帰係数の

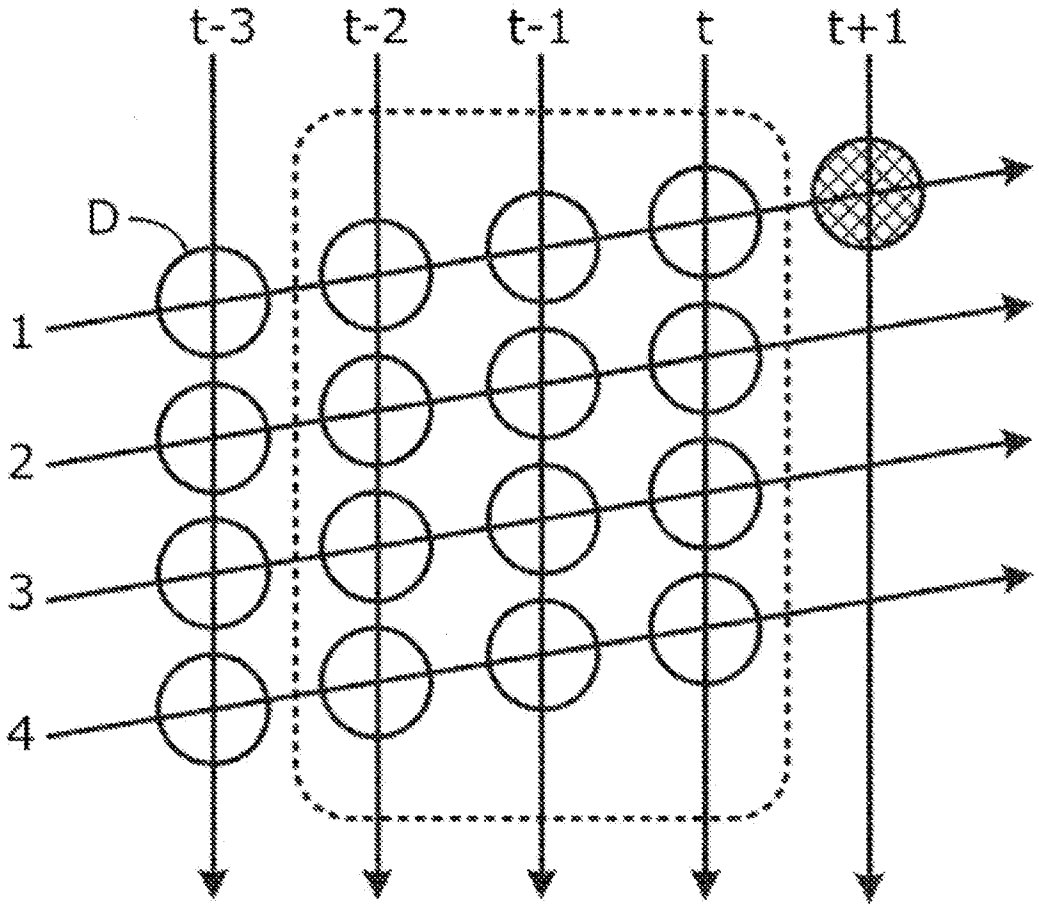
うち絶対値が閾値以上の回帰係数に対応するデータを選択するための選択情報を送信し、

前記第１算出法と異なる第２算出法を用いて、受信した前記選択データを使用して回帰係数を算出し、

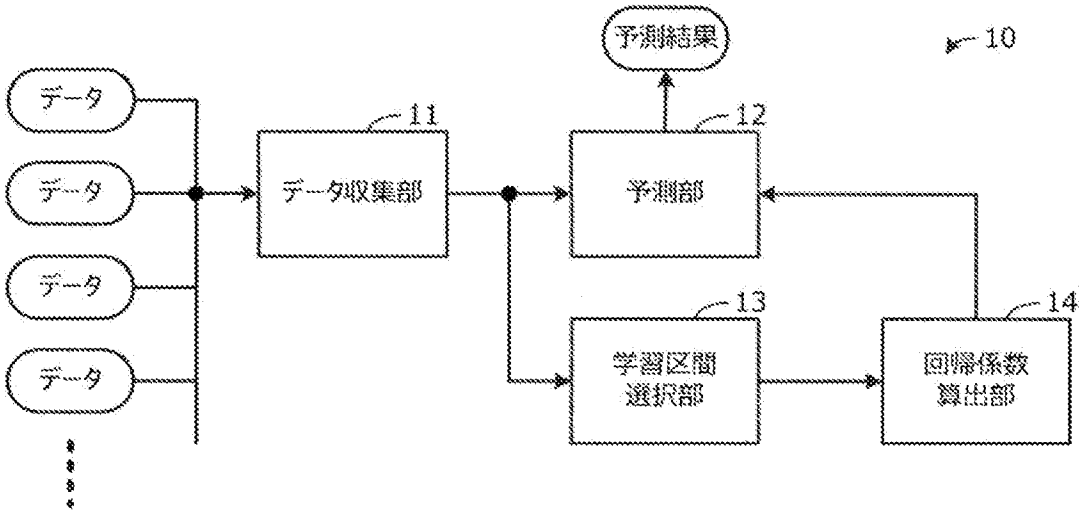
受信した前記選択データ、及び前記第２算出法を用いて算出された回帰係数に基づいて予測した予測結果を出力する、

データ予測方法。

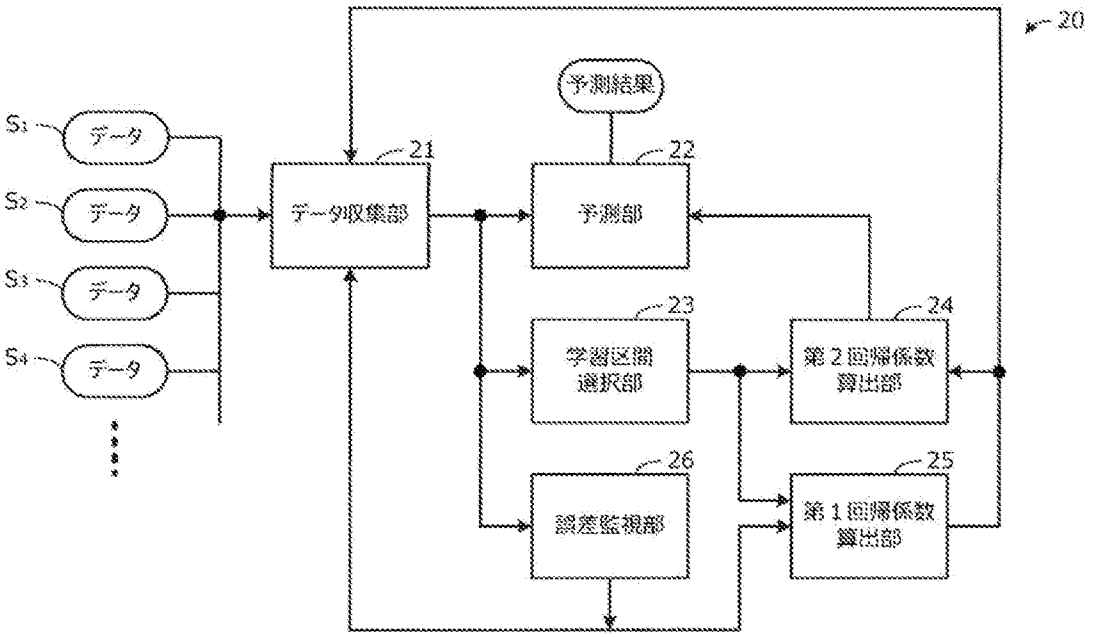
[図1]


t-3
t-2
t-1
t
t+1
D
1
2
3
4

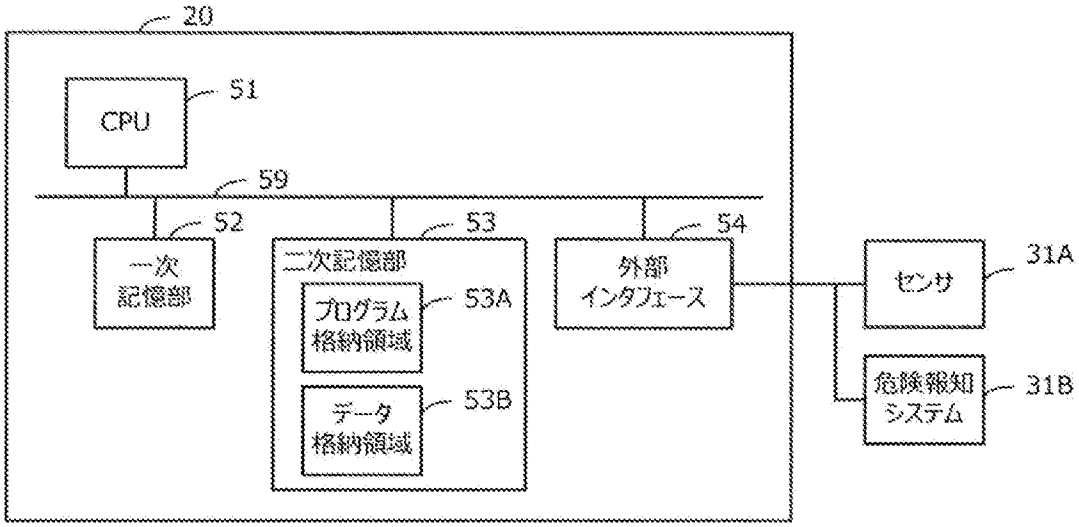
[図2]


予測結果
10
データ
データ
データ
データ
11
データ収集部
12
予測部
13
学習区間
選択部
14
回帰係数
算出部

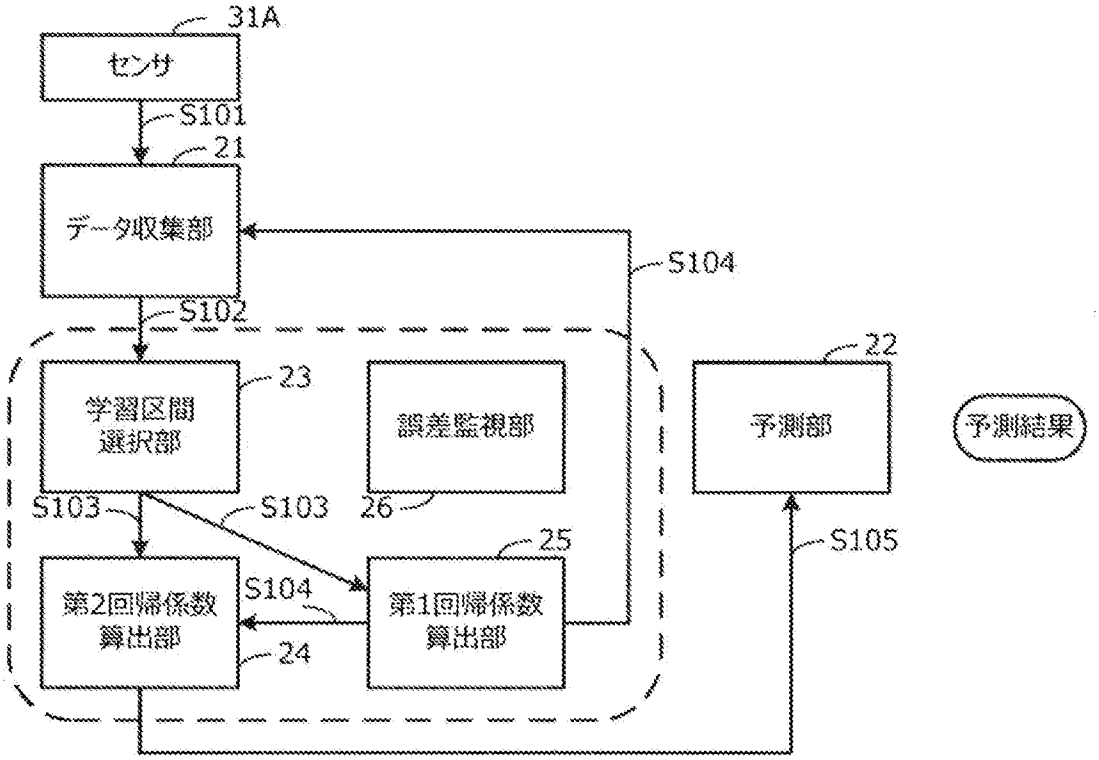
[図3]


20
予測結果
S1
データ
S2
データ
S3
データ
S4
データ
21
データ収集部
22
予測部
23
学習区間
選択部
24
第2回帰係数
算出部
26
誤差監視部
25
第1回帰係数
算出部

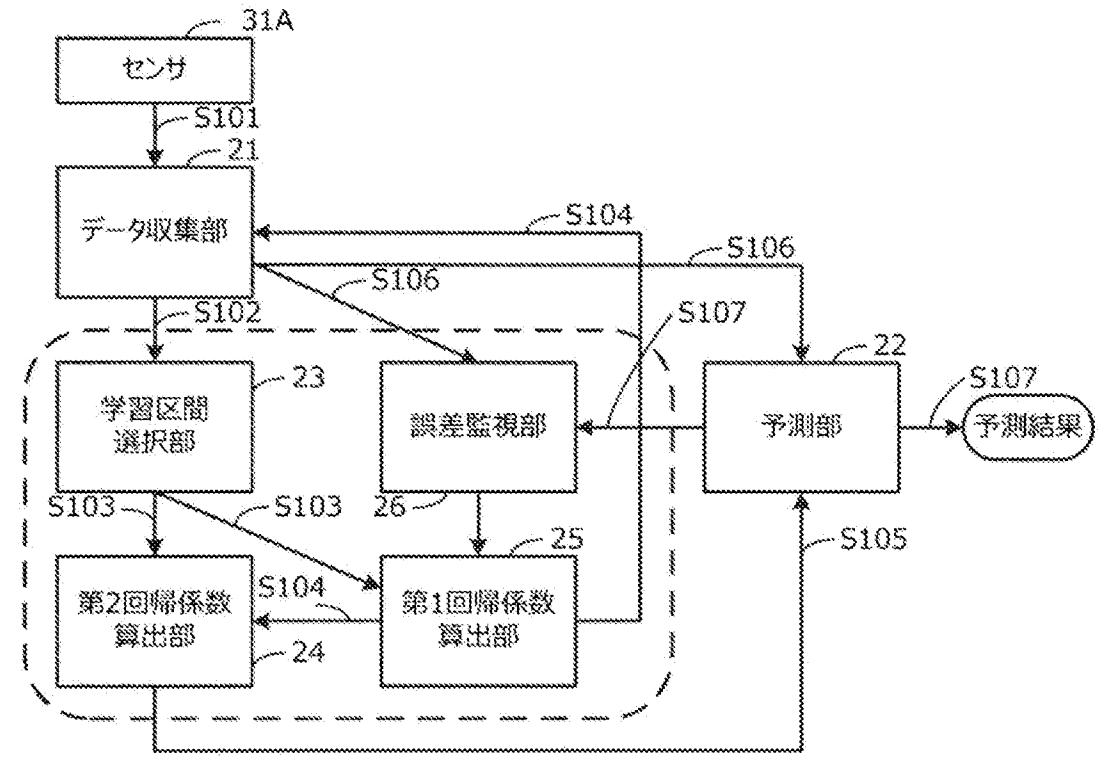
[図4]


20
CPU
51
59
52
一次
記憶部
53
二次記憶部
プログラム
格納領域
53A
データ
格納領域
53B
54
外部
インタフェース
センサ
31A
危険報知
システム
31B

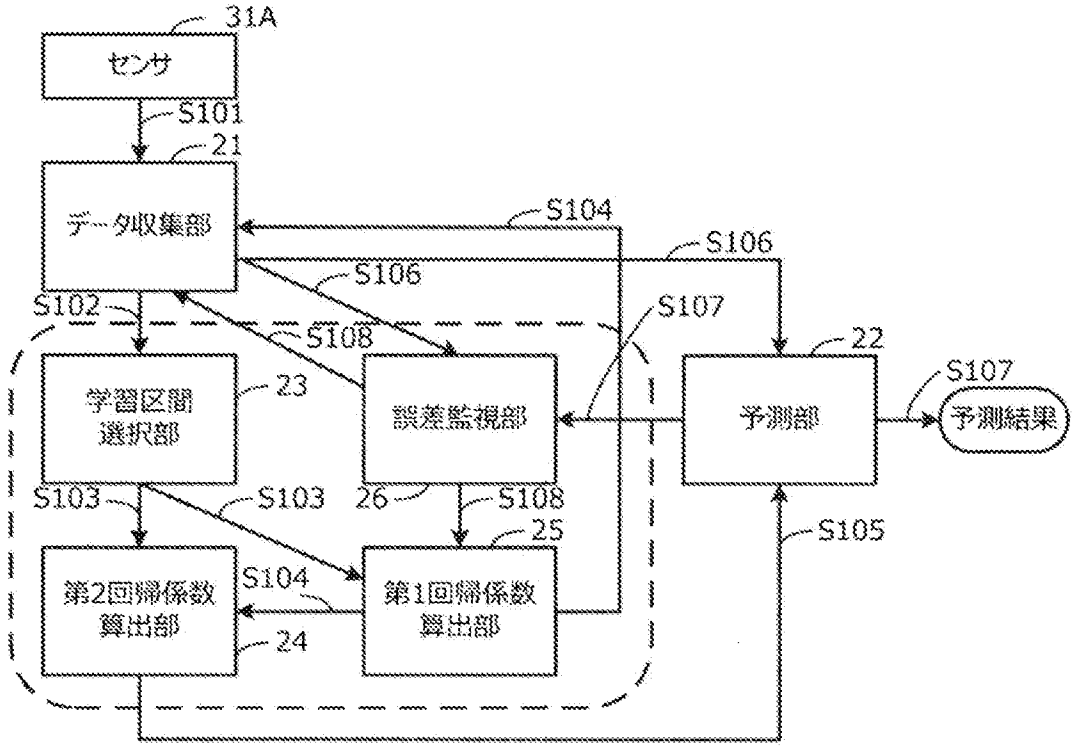
[図5]


31A
センサ
S101
21
データ収集部
S104
S102
22
23
学習区間
選択部
誤差監視部
予測部
予測結果
S103
S103
26
25
S105
第2回帰係数
算出部
S104
第1回帰係数
算出部
24

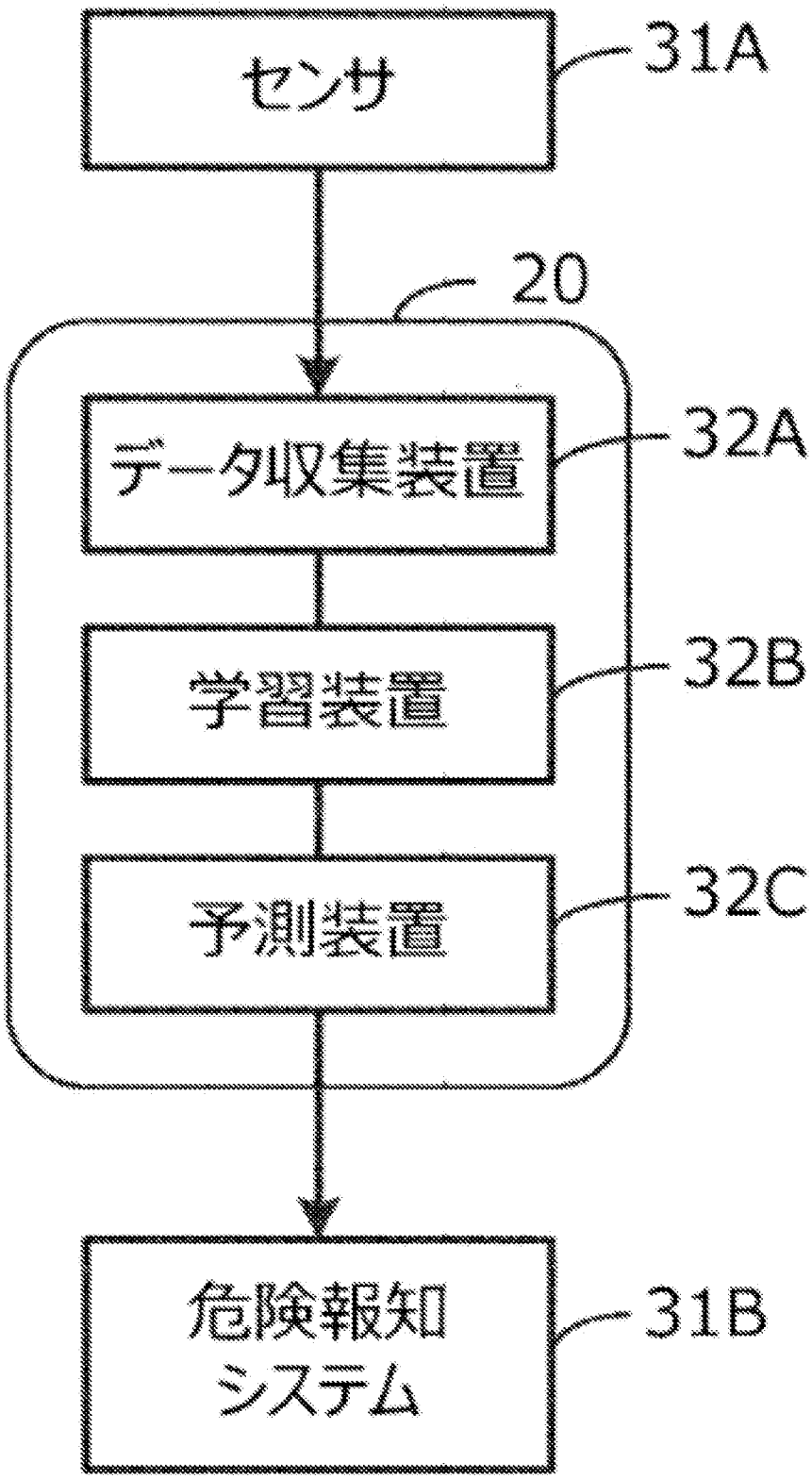
[図6]


31A
センサ
S101
21
データ収集部
S104
S106
S106
S107
S102
22
S107
23
学習区間
選択部
誤差監視部
予測部
予測結果
S103
S103
26
25
S105
第2回帰係数
算出部
S104
第1回帰係数
算出部
24

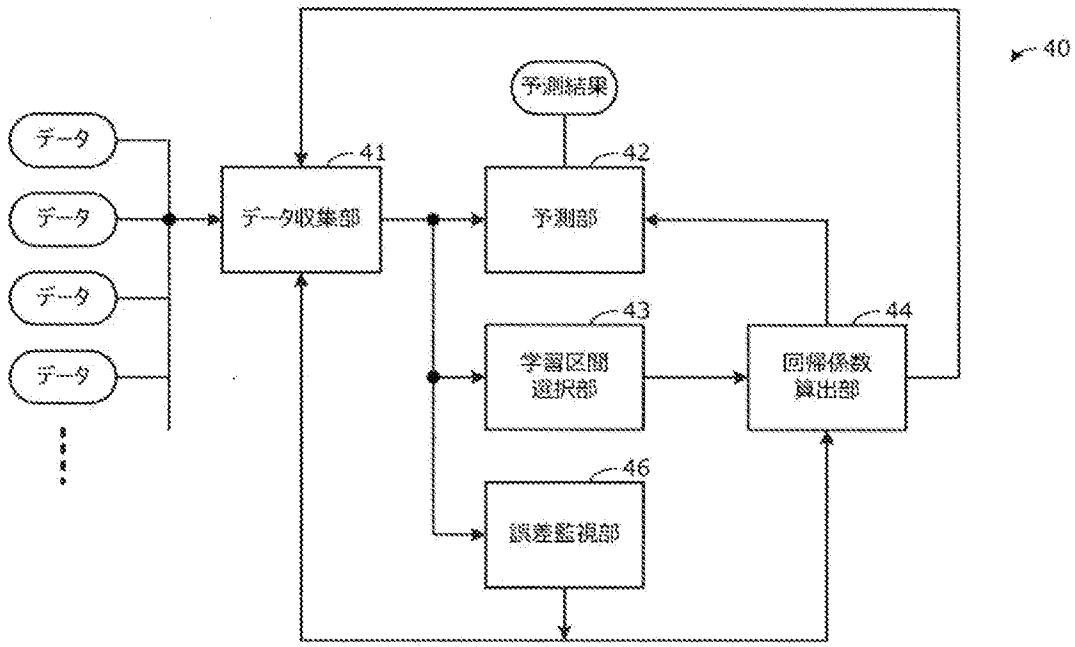
[図7]


31A
センサ
S101
21
データ収集部
S104
S106
S106
S102
S107
S108
22
S107
学習区間
選択部
23
誤差監視部
予測部
予測結果
S103
S103
26
S108
25
S105
第2回帰係数
算出部
S104
第1回帰係数
算出部
24

[図8]



[図9]


40
予測結果
データ
データ
データ
データ
41
データ収集部
42
予測部
43
学習区間
選択部
44
回帰係数
算出部
46
誤差監視部

INTERNATIONAL SEARCH REPORT

International application No.
PCT/JP2020/010304

A. CLASSIFICATION OF SUBJECT MATTER
G06N 20/00(2019.01)i
FI: G06N20/00 130

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
G06N20/00

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Published examined utility model applications of Japan	1922-1996
Published unexamined utility model applications of Japan	1971-2020
Registered utility model specifications of Japan	1996-2020
Published registered utility model applications of Japan	1994-2020

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X A	JP 2019-32185 A (TOSHIBA CORP.) 28.02.2019 (2019-02-28) paragraphs [0009]-[0100], fig. 1-11	4-5 1-3, 6-7
A	JP 2018-151883 A (TOSHIBA CORP.) 27.09.2018 (2018-09-27) paragraph [0020]	1-7
A	JP 2018-109876 A (TOSHIBA CORP.) 12.07.2018 (2018-07-12) paragraphs [0068]-[0080]	1-7
A	JP 2018-10477 A (FUJITSU LTD.) 18.01.2018 (2018-01-18) entire text, all drawings	1-7
A	JP 2017-130041 A (FUJITSU LTD.) 27.07.2017 (2017-07-27) entire text, all drawings	1-7

☐ Further documents are listed in the continuation of Box C.

☒ See patent family annex.

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier application or patent but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
28 April 2020 (28.04.2020)	19 May 2020 (19.05.2020)

Name and mailing address of the ISA/	Authorized officer
Japan Patent Office 3-4-3, Kasumigaseki, Chiyoda-ku, Tokyo 100-8915, Japan	Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/JP2020/010304

Patent Documents referred in the Report	Publication Date	Patent Family	Publication Date
JP 2019-32185 A	28 Feb. 2019	US 2019/0041235 A1 paragraphs [0023]-[0116], fig. 1-11	
JP 2018-151883 A	27 Sep. 2018	US 2018/0260726 A1 paragraph [0036] KR 10-2018-0104542 A	
JP 2018-109876 A	12 Jul. 2018	US 2018/0189242 A1 paragraphs [0090]-[0102] EP 3346428 A1	
JP 2018-10477 A	18 Jan. 2018	(Family: none)	
JP 2017-130041 A	27 Jul. 2017	(Family: none)	

国際調査報告

国際出願番号
PCT/JP2020/010304

A. 発明の属する分野の分類（国際特許分類（ＩＰＣ））

G06N 20/00(2019.01)i
FI: G06N20/00 130

B. 調査を行った分野

調査を行った最小限資料（国際特許分類（ＩＰＣ））

G06N20/00

最小限資料以外の資料で調査を行った分野に含まれるもの

日本国実用新案公報	１９２２－１９９６年
日本国公開実用新案公報	１９７１－２０２０年
日本国実用新案登録公報	１９９６－２０２０年
日本国登録実用新案公報	１９９４－２０２０年

国際調査で使用した電子データベース（データベースの名称、調査に使用した用語）

C. 関連すると認められる文献

引用文献のカテゴリー*	引用文献名　及び一部の箇所が関連するときは、その関連する箇所の表示	関連する請求項の番号
X	JP 2019-32185 A (株式会社東芝) 28.02.2019 (2019-02-28) [0009]-[0100],図1-11	4-5
A		1-3,6-7
A	JP 2018-151883 A (株式会社東芝) 27.09.2018 (2018-09-27) [0020]	1-7
A	JP 2018-109876 A (株式会社東芝) 12.07.2018 (2018-07-12) [0068]-[0080]	1-7
A	JP 2018-10477 A (富士通株式会社) 18.01.2018 (2018-01-18) 全文、全図	1-7
A	JP 2017-130041 A (富士通株式会社) 27.07.2017 (2017-07-27) 全文、全図	1-7

☐ C欄の続きにも文献が列挙されている。

☑ パテントファミリーに関する別紙を参照。

* 引用文献のカテゴリー
"A" 特に関連のある文献ではなく、一般的技術水準を示すもの
"E" 国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの
"L" 優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献（理由を付す）
"O" 口頭による開示、使用、展示等に言及する文献
"P" 国際出願日前で、かつ優先権の主張の基礎となる出願の日の後に公表された文献
"T" 国際出願日又は優先日後に公表された文献であって出願と抵触するものではなく、発明の原理又は理論の理解のために引用するもの
"X" 特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの
"Y" 特に関連のある文献であって、当該文献と他の１以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの
"&" 同一パテントファミリー文献

国際調査を完了した日	国際調査報告の発送日
28.04.2020	19.05.2020

名称及びあて先
日本国特許庁(ISA/JP)
〒100-8915
日本国
東京都千代田区霞が関三丁目４番３号

権限のある職員（特許庁審査官）
山本　俊介 5B 5087
電話番号 03-3581-1101 内線 3545

国際調査報告
パテントファミリーに関する情報

国際出願番号
PCT/JP2020/010304

引用文献	公表日	パテントファミリー文献	公表日
JP 2019-32185 A	28.02.2019	US 2019/0041235 A1 [0023]-[0116],FIGs.1-11	
JP 2018-151883 A	27.09.2018	US 2018/0260726 A1 [0036] KR 10-2018-0104542 A	
JP 2018-109876 A	12.07.2018	US 2018/0189242 A1 [0090]-[0102] EP 3346428 A1	
JP 2018-10477 A	18.01.2018	（ファミリーなし）	
JP 2017-130041 A	27.07.2017	（ファミリーなし）	