

(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
2022年4月14日(14.04.2022)



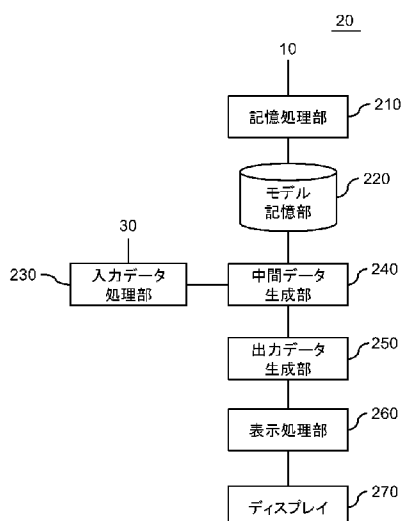
(10) 国際公開番号

WO 2022/075215 A1

- (51) 国際特許分類:
G06N 20/00 (2019.01)
- (21) 国際出願番号: PCT/JP2021/036407
- (22) 国際出願日: 2021年10月1日(01.10.2021)
- (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 2020-171188 2020年10月9日(09.10.2020) JP
- (71) 出願人: 株式会社エンビジョン A E S C
ジャパン(ENVISION AESC JAPAN LTD.) [JP/
JP]; 〒2520012 神奈川県座間市広野台二丁
目10番1号 Kanagawa (JP).
- (72) 発明者: 陳 九廷(CHEN Jiuting); 〒2520012 神
奈川県座間市広野台二丁目10番1号 株
式会社エンビジョン A E S C ジャパン
内 Kanagawa (JP). 嶋脇 秀徳(SHIMAWAKI
Hidenori); 〒2520012 神奈川県座間市広野台
- 二丁目10番1号 株式会社エンビジョン
A E S C ジャパン内 Kanagawa (JP). 林 逸郎
(HAYASHI Itsuro); 〒2520012 神奈川県座間市
広野台二丁目10番1号 株式会社エンビジョ
ン A E S C ジャパン内 Kanagawa (JP).
- (74) 代理人: 速水 進治, 外(HAYAMI Shinji et al.);
〒1410031 東京都品川区西五反田7丁目9番2
号 K D X 五反田ビル9階 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, IT, 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 PROCESSING SYSTEM, MODEL GENERATION DEVICE, DATA PROCESSING METHOD, MODEL GENERATION METHOD, AND PROGRAM

(54) 発明の名称: データ処理システム、モデル生成装置、データ処理方法、モデル生成方法、及びプログラム



210 ... Storage processing unit
220 ... Model storage unit
230 ... Input data processing unit
240 ... Intermediate data generation unit
250 ... Output data generation unit
260 ... Display processing unit
270 ... Display

(57) Abstract: A data processing system (20) is provided with an input data processing unit (230), an intermediate data generation unit (240), and an output data generation unit (250). The input data processing unit (230) acquires input data and converts the input data in accordance with a conversion rule. An example of this conversion rule is to increase the number of values included in the input data by adding dummy data to the input data. The intermediate data generation unit (240) generates intermediate data by processing the converted input data. This intermediate data is, for example, data in a plurality of rows and/or a plurality of columns that comprise a plurality of values. The intermediate data includes a meaningful value (that is to say, value to be obtained) and a value to be dummy. The output data generation unit (250) generates the output data by selecting the meaningful value from the intermediate data generated by the intermediate data generation unit (240).



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: データ処理システム (20) は、入力データ処理部 (230)、中間データ生成部 (240)、及び出力データ生成部 (250) を備えている。入力データ処理部 (230) は、入力データを取得し、当該入力データを変換ルールに従って変換する。この変換ルールの一例は、入力データにダミーデータを追加することにより、入力データに含まれる値の数を増やすことである。中間データ生成部 (240) は、変換後の入力データを処理することにより、中間データを生成する。この中間データは、例えば複数の値からなる複数行及び／または複数列のデータである。中間データは、意味がある値 (すなわち求めたい値) と、ダミーとなる値とを含んでいる。出力データ生成部 (250) は、中間データ生成部 (240) が生成した中間データのうち意味がある値を選択することにより出力データを生成する。

明 細 書

発明の名称：

データ処理システム、モデル生成装置、データ処理方法、モデル生成方法、及びプログラム

技術分野

[0001] 本発明は、データ処理システム、モデル生成装置、データ処理方法、モデル生成方法、及びプログラムに関する。

背景技術

[0002] 近年は、機械学習によって生成されたモデルが様々な用途で使用されている。例えば特許文献1には、二次電池の状態を推定する際にニューラルネットワークを用いた学習を用いることが記載されている。

先行技術文献

特許文献

[0003] 特許文献1：特開2008-232758号公報

発明の概要

発明が解決しようとする課題

[0004] データの解析に用いられるモデル、例えば機械学習によって生成されたモデルの価値は高い。一方、このモデルの用途が広がると、このモデルを用いた装置がネットワークに接続する場合のように、第三者がモデルに接続できるケースが増えてくる。この場合、このモデルに入力データを送信し、このモデルによる出力データを受信することも可能になる。すると、第三者は、入力データの出力データの組み合わせを複数得ることができるため、モデルの構造を推定できる可能性が出てくる。

[0005] 本発明の目的の一例は、データの解析に用いられるモデルの秘匿性を高めることにある。

課題を解決するための手段

[0006] 本発明によれば、複数の値からなる入力データを取得し、当該入力データを変換ルールに従って変換する入力データ処理部と、

変換後の前記入力データを処理することにより、複数の値からなる複数行及び／または複数列のデータからなる中間データを出力する中間データ生成部と、

前記中間データに所定の演算を行った後に予め定められた位置にある少なくとも一つの値を選択することにより出力データを生成する出力データ生成部と、

を備え、

前記中間データ生成部と、前記出力データ生成部とは異なる装置に設けられているデータ処理システムが提供される。

[0007] 本発明によれば、請求項 1 に記載のデータ処理システムが有する中間データ生成部で用いられるモデルを生成するモデル生成装置であって、

正常なデータからなる第 1 訓練データ、及び一部に異常なデータを含む第 2 訓練データを取得する訓練データ取得部と、

前記モデルを、前記第 1 訓練データを用いて訓練し、その後に、前記変換ルールによって変換された後の前記第 2 訓練データを用いて訓練する工程を、複数回繰り返すモデル生成部と、

を備えるモデル生成装置が提供される。

[0008] 本発明によれば、コンピュータが、

複数の値からなる入力データを取得し、当該入力データを変換ルールに従って変換する入力データ処理と、

変換後の前記入力データを処理することにより、複数の値からなる複数行及び／または複数列のデータからなる中間データを出力する中間データ生成処理と、

前記中間データに所定の演算を行った後に予め定められた位置にある少なくとも一つの値を選択することにより出力データを生成する出力データ生成処理と、

を行い、

前記中間データ生成処理と、前記出力データ生成処理とが、異なるコンピュータによって行われるデータ処理方法が提供される。

[0009] 本発明によれば、上記データ処理システムが有する中間データ生成部で用いられるモデルを生成するモデル生成方法であって、

コンピュータに、

正常なデータからなる第1訓練データ、及び一部に異常なデータを含む第2訓練データを取得する訓練データ取得処理と、

前記モデルを、前記第1訓練データを用いて訓練し、その後に、前記変換ルールによって変換された後の前記第2訓練データを用いて訓練する工程を、複数回繰り返すモデル生成処理と、
を行わせるモデル生成方法が提供される。

[0010] 本発明によれば、コンピュータに、

複数の値からなる入力データを取得し、当該入力データを変換ルールに従って変換する入力データ処理機能と、

変換後の前記入力データを処理することにより、複数の値からなる複数行及び／または複数列のデータからなる中間データを出力する中間データ生成機能と、

を持たせるプログラムが提供される。

[0011] 本発明によれば、コンピュータに、

上記した中間データ生成機能によって生成された前記中間データに所定の演算を行った後に予め定められた位置にある少なくとも一つの値を選択することにより出力データを生成する出力データ生成機能と、
を持たせるプログラムが提供される。

[0012] 本発明によれば、コンピュータに、請求項1に記載のデータ処理システムが有する中間データ生成部で用いられるモデルを生成する機能を持たせるプログラムであって、

前記コンピュータに、

正常なデータからなる第1訓練データ、及び一部に異常なデータを含む第2訓練データを取得する訓練データ取得機能と、

前記モデルを、前記第1訓練データを用いて訓練し、その後に、前記変換ルールによって変換された後の前記第2訓練データを用いて訓練する工程を、複数回繰り返すモデル生成機能と、
を持たせるプログラムが提供される。

発明の効果

[0013] 本発明によれば、データの解析に用いられるモデルの秘匿性は高まる。

図面の簡単な説明

[0014] [図1]実施形態に係るモデル生成装置及びデータ処理システムの使用環境の一例を説明するための図である。

[図2]モデル生成装置の機能構成の一例を示す図である。

[図3]データ処理システムの機能構成の一例を示す図である。

[図4]モデル生成装置の事前処理部及びデータ処理システムの入力データ処理部が用いる変換ルールの第1例を説明するための図である。

[図5]データ処理システムの入力データ処理部が用いる変換ルールの第2例を説明するための図である。

[図6]モデル生成装置の事前処理部及びデータ処理システムの入力データ処理部が用いる変換ルールの第3例を説明するための図である。

[図7]モデル生成装置のハードウェア構成例を示す図である。

[図8]モデル生成装置が行うモデルの生成処理の第1例を示すフローチャートである。

[図9]モデル生成装置が行うモデルの生成処理の第2例を示すフローチャートである。

[図10]データ処理システムが行う処理の一例を示すフローチャートである。

[図11]仮モデルの生成方法を説明するための図である。

発明を実施するための形態

[0015] 以下、本発明の実施の形態について、図面を用いて説明する。尚、すべて

の図面において、同様な構成要素には同様の符号を付し、適宜説明を省略する。

[0016] 図1は、実施形態に係るモデル生成装置10及びデータ処理システム20の使用環境の一例を説明するための図である。データ処理システム20は、モデル生成装置10が生成したモデルを用いて入力データを処理し、出力データを出力する。入力データは、複数の値からなる。これら複数の値の一例は、対象物の状態を互いに異なる指標で測定した結果である。以下の説明において、対象物は蓄電池（二次電池）30であり、入力データは蓄電池30の状態の測定結果である。この場合、入力データは、例えば蓄電池30の出力電圧、出力電流値、及び温度を含んでいる。ただし、対象物及び入力データはこれらに限定されない。

[0017] 本図に示す例において、モデル生成装置10及びデータ処理システム20は、蓄電池30と共に使用される。データ処理システム20の一部の機能は、蓄電池30のBMS（Battery Management System）であってもよい。ただし、後述する中間データ生成部240と出力データ生成部250は、互いに異なる装置に設けられている。そして一つのデータ処理システム20が複数の蓄電池30に接続しており、これら複数の蓄電池30に関する処理を行う。

[0018] 蓄電池30は機器40に電力を供給する。一例として、機器40は、例えば電動車両などの車両である。ただし、蓄電池30が家庭用の蓄電池の場合、機器40は家庭で用いられる電気機器となる。この場合、蓄電池30は機器40の外部に位置する。また蓄電池30は、系統電力網に接続していてもよい。この場合、蓄電池30は、供給される電力を平準化するために用いられる。具体的には、機器40は、電力が余っている時には電力を蓄え、電力が不測している時には電力を供給する。

[0019] データ処理システム20は、蓄電池30の状態を推定する。ここで推定される状態は、例えば、蓄電池30の残容量（単位はAh）、充電率（SOC：State Of Charge）、及びSOH（State Of Health）の少なくとも一つで

ある。SOHは、例えば「現在の満充電容量（Ah）／初期の満充電容量（Ah）×100（％）」である。なお、蓄電池30の状態はこれらに限定されない。この推定において、上記したように、データ処理システム20はモデルを用いる。モデル生成装置10は、データ処理システム20が用いるモデルの少なくとも一つを、機械学習、例えばニューラルネットワークを用いて生成し、かつ更新する。

[0020] モデル生成装置10は、複数の蓄電池30から蓄電池30の状態に関するデータの測定値（以下、実績データと記載）を取得してもよい。この場合、複数の実績データの一部は、機械学習の訓練データとして用いられ、実績データの残りの少なくとも一部は、モデルを検証するために用いられてもよい。

[0021] 実績データは、蓄電池30の種類（例えば品名や型番）を特定する情報を含んでいるのが好ましい。このようにすると、モデル生成装置10は、蓄電池30の種類別にモデルを生成することができる。そしてデータ処理システム20は、当該データ処理システム20が接続している蓄電池30の種類に対応するモデルをモデル生成装置10から取得し、使用することができる。したがって、データ処理システム20による蓄電池30の状態の推定精度は高くなる。

[0022] なお、実績データの少なくとも一部は、実績データを集めることを主目的として使用されている蓄電池30によって生成されてもよい。

[0023] 図2は、モデル生成装置10の機能構成の一例を示す図である。本図に示す例において、モデル生成装置10は、訓練データ取得部120及びモデル生成部140を備えている。訓練データ取得部120は、複数の訓練データを取得する。一例として、訓練データのそれぞれは、蓄電池のある充放電サイクルにおける状態を示す指標（例えば電流、電圧、及び温度）を含む訓練用測定データを入力データとして、当該蓄電池の性能を示す値（例えば残容量、SOC、及びSOHの少なくとも一つ）である訓練用出力データを目標値としている。モデル生成部140は、複数の訓練データを機械学習するこ

とにより、モデルを生成する。

[0024] モデル生成部 140 は、複数の機械学習アルゴリズム（例えば L S T M（Long Short-Term Memory）、D N N（Deep Neural Network）、及び L R（Linear Regression）など）を用いて、複数のモデルを生成してもよい。

[0025] 本図に示す例において、訓練データ取得部 120 は訓練データ記憶部 110 から訓練データを取得する。訓練データ記憶部 110 は、モデル生成装置 10 の一部であってもよいし、モデル生成装置 10 の外部に設けられていてもよい。訓練データ記憶部 110 が記憶している訓練データは、正常なデータからなる第 1 訓練データ、及び一部に異常なデータを含む第 2 訓練データを含んでいる。第 1 訓練データの一例は、上記した実績データである。第 2 訓練データの一例は、実績データとしては存在し得ない値を含んでおり、例えば人またはコンピュータによって生成される。

[0026] モデル生成装置 10 は、さらに事前処理部 130 を備えている。事前処理部 130 は、少なくとも第 2 訓練データを、予め定められた変換ルールに従って変換する。モデル生成部 140 は、変換後の第 2 訓練データを用いてモデルを生成する。ここで事前処理部 130 は、さらに第 1 訓練データを、予め定められた変換ルールに従って変換してもよい。この場合、モデル生成部 140 は、変換後の第 1 訓練データを用いてモデルを生成する。

[0027] 上記した変換ルールの一例は、訓練データの入力データに対してダミーデータを追加することにより、当該入力データを構成する値の数を増やすとともに、訓練データの出力データに対してもダミーデータを追加することにより、当該出力データを構成する値の数を増やす処理である。例えば入力データが 3 つの値からなる場合、変換後の入力データは、当初から入力データに含まれていた 3 つの値と、少なくとも一つの値からなるダミーデータと、を含んでいるため、4 つ以上の値から構成されることになる。なお、この変換ルールの詳細例については後述する。

[0028] モデル生成部 140 が生成したモデルは、モデル記憶部 150 に記憶される。そしてモデル記憶部 150 に記憶されたモデルは、モデル送信部 160

によってデータ処理システム 20 に送信される。本図に示す例において、モデル記憶部 150 及びモデル送信部 160 は、モデル生成装置 10 の一部となっている。ただし、モデル記憶部 150 及びモデル送信部 160 の少なくとも一方はモデル生成装置 10 の外部の装置になっていてもよい。

[0029] 図 3 は、データ処理システム 20 の機能構成の一例を示す図である。データ処理システム 20 は、入力データ処理部 230、中間データ生成部 240、及び出力データ生成部 250 を備えている。

[0030] 中間データ生成部 240 と出力データ生成部 250 は、互いに異なる装置となっている。一例として、データ処理システム 20 のうち出力データ生成部 250 がクラウドサーバ上に設けられており、データ処理システム 20 の他の機能が端末上に設けられている。ただし、後述する記憶処理部 210、モデル記憶部 220、及び表示処理部 260 もクラウドサーバ上に設けられていてもよい。

[0031] 入力データ処理部 230 は、入力データを取得し、当該入力データを変換ルールに従って変換する。入力データの構成は、モデル生成装置 10 が用いる訓練データの入力データと同じである。例えば入力データ処理部 230 は、蓄電池 30 の状態を検出するセンサ（例えば電流計、電圧計、及び温度計）から入力データを取得する。また入力データ処理部 230 が用いる変換ルールは、モデル生成装置 10 の事前処理部 130 が用いる変換ルールと同じである。この変換ルールの一例は、上記したように、入力データにダミーデータを追加することにより、入力データに含まれる値の数を増やすことである。

[0032] 中間データ生成部 240 は、変換後の入力データを処理することにより、中間データを生成する。この中間データは、例えば複数の値からなる複数行及び／または複数列のデータである。中間データ生成部 240 は、モデル生成装置 10 が生成したモデルを用いて中間データを生成する。上記したように、モデル生成装置 10 はモデルを生成する際に、訓練データの出力データにダミーデータを追加している。このため、生成される中間データは、意味

がある値（すなわち求めたい値）と、ダミーとなる値とを含んでいる。

[0033] 出力データ生成部 250 は、中間データ生成部 240 が生成した中間データに所定の演算を行った後に予め定められた位置にある少なくとも一つの値（すなわち中間データのうち意味がある値）を選択することにより出力データを生成する。ここで行われる所定の演算は、例えば加算であるが、減算、乗算、又は除算であってもよいし、加算、減算、乗算、及び除算の少なくとも2つを組み合わせた演算であってもよい。

[0034] 本図に示す例において、データ処理システム 20 は、さらに記憶処理部 210、表示処理部 260、及びディスプレイ 270 を備えている。

[0035] 記憶処理部 210 は、モデル生成装置 10 からモデルを取得し、モデル記憶部 220 に記憶させる。記憶処理部 210 は、モデル生成装置 10 からモデルを更新するためのデータ（例えばモデルのパラメータ）を取得した場合、このデータを用いて、モデル記憶部 220 に記憶されているモデルを更新する。この更新処理は、繰り返し行われるのが好ましい。本図に示す例において、モデル記憶部 220 はデータ処理システム 20 の一部となっている。ただし、モデル記憶部 220 はデータ処理システム 20 の外部の装置であってもよい。

[0036] 表示処理部 260 は、出力データ生成部 250 が生成した出力データをディスプレイ 270 に表示させる。ディスプレイ 270 は、機器 40 の使用者が視認可能な位置に配置されている。例えば機器 40 が車両の場合、ディスプレイ 270 は当該車両の内部（例えば運転席の前又は斜め前）に設けられている。

[0037] 図 4 は、モデル生成装置 10 の事前処理部 130 及びデータ処理システム 20 の入力データ処理部 230 が用いる変換ルールの第 1 例を説明するための図である。図 4（A）は入力データの第 1 例を示しており、図 4（B）は図 4（A）に対応する出力データの一例を示している。

[0038] 図 4（A）に示す例において、入力データは行データであり、複数の値（例えば [1, 1, 1] という 3 つの値）によって構成されている。そして変

換ルールは、この行の所定の位置にダミーデータを挿入することである。その結果、入力データは、2列目にダミーの値である「0」が追加されている。そして2列目以降の値は一つずつ繰り下がっている。

[0039] そして、図4（B）に示す例において、本来の出力データは行データであり、所定の数の値（例えば[1, 3, 1]という3つの値）によって構成されている。これに対し、モデル生成装置10が生成したモデルは、変換後の入力データを用いることを前提としており、その結果、本来の出力データよりも多い数の値（例えば[-3, 1, 3, 1, -2]という5つの値）からなる行データとなる。ただしこの行データは、予め定められた位置（すなわち予め定められた列）に、本来の出力データ（[1, 3, 1]）を含んでいるとともに、残りの位置（列）にダミーの値を含んでいる。出力データ生成部250は[-3, 1, 3, 1, -2]から、[1, 3, 1]を抽出する処理を行う。なお、本来の出力データとなる複数の値は連続している必要はない。

[0040] なお、モデル生成装置10の事前処理部130は、訓練データに含まれる入力データ（例えば[1, 1, 1]）に対して所定の値のダミーデータ（例えば0）を所定の位置（例えば2列目）に追加する。また事前処理部130は、訓練データに含まれる出力データ（例えば[1, 3, 1]）に対して所定のダミーデータを所定の位置（例えば1列目と5列目）を追加する。ここで出力データに追加される所定のダミーデータの値は、本来の出力データに含まれる値が取り得る範囲内にあるのが好ましい。このようにすると、中間データ生成部240が生成した中間データに含まれるダミーの値（すなわち出力データとして選択されない値）は、本来の出力データとなる複数の値が取り得る範囲内になる。したがって、第三者が中間データを見たとしても、この第三者は中間データに含まれる複数の値のうちどれが意味のあるデータ（すなわち本来の出力データ）であるか、判断できない。このため、第三者は入力データと出力データの組み合わせを特定できなくなり、その結果、モデル生成装置10が生成したモデルの秘匿性は高くなる。

[0041] なお、モデルをこのようにするためには、モデル生成装置 10 が用いる第 2 訓練データにおいて、出力データに含まれるダミーの値を、本来の出力データとなる複数の値が取り得る範囲内にすればよい。

[0042] 図 5 は、データ処理システム 20 の入力データ処理部 230 が用いる変換ルール第 2 例を説明するための図である。図 5 (A) は入力データの第 1 例を示しており、図 5 (B) は図 5 (A) に対応する出力データの一例を示している。

[0043] 本図に示す処理は、図 4 を用いて説明した処理とほぼ同様である。ただし、変換ルールは、さらに、「入力データに所定の条件を満たす値が含まれていたときに、当該値を他の値に置換すること」を含んでいる。第三者がモデルの構造を推定することを目的として、コンピュータを用いて多数の入力データをデータ処理システム 20 に入力する場合、入力データは特定の値（例えば 0 や 1 などの整数）を含むことが多い。そこで変換ルールにおける「所定の条件を満たす値」として、コンピュータが自動でデータを生成するときに用いやすい値を設定しておく、変換ルールによってこの値は他の値に置き換えられる。従って、モデルの秘匿性は高くなる。

[0044] ここで、「所定の条件を満たす値」と、置き換えた後の値である「他の値」との差を小さくすると、入力データが「所定の条件を満たす値」を含んでいた場合であっても、出力データ生成部 250 が出力する出力データの値と、値を置き換えなかった場合の出力データ（本来の出力データの値）との差は小さくなる。このため、変換ルールの値の置換に起因した誤差は小さくなる。

[0045] なお、上記した値の置き換えは、モデル生成装置 10 の事前処理部 130 では行われなくてもよい。

[0046] 図 6 は、モデル生成装置 10 の事前処理部 130 及びデータ処理システム 20 の入力データ処理部 230 が用いる変換ルール第 3 例を説明するための図である。

[0047] 本図に示す処理は、図 4 又は図 5 を用いて説明した処理を前提にしている

。このため、図6（A）に示す処理は、図5（A）（又は図4（A））に示した処理と同じである。その上で、中間データ生成部240が生成した中間データは、これらのデータを構成する少なくとも一つの値（ダミーとなる値ではなく本来のデータを構成する値）に所定の演算を行うことにより、本来のデータが得られるようになっている。この演算は、出力データ生成部250が行う。例えば図6（B）に示す例では、3列目の値に定数（例えば1）を加えることにより、本来のデータが得られるようになっている。

[0048] このようにするためには、例えばモデル生成部140が用いる訓練データにおいて、出力データに対して、中間データ生成部240（又は出力データ生成部250）が行う演算とは逆の演算（例えば3列目の値から定数（例えば1）を引く）を行い、この逆演算を行った後の訓練データを用いて、中間データ生成部240が用いるモデルを生成すればよい。

[0049] 本図に示す例によれば、モデルが出力する中間データのうち本来のデータを示す値は、本来の値とは異なる値になる。このため、モデルの秘匿性はさらに高くなる。

[0050] 図7は、モデル生成装置10のハードウェア構成例を示す図である。モデル生成装置10は、バス1010、プロセッサ1020、メモリ1030、ストレージデバイス1040、入出力インタフェース1050、及びネットワークインタフェース1060を有する。

[0051] バス1010は、プロセッサ1020、メモリ1030、ストレージデバイス1040、入出力インタフェース1050、及びネットワークインタフェース1060が、相互にデータを送受信するためのデータ伝送路である。ただし、プロセッサ1020などを互いに接続する方法は、バス接続に限定されない。

[0052] プロセッサ1020は、CPU（Central Processing Unit）やGPU（Graphics Processing Unit）などで実現されるプロセッサである。

[0053] メモリ1030は、RAM（Random Access Memory）などで実現される主記憶装置である。

- [0054] ストレージデバイス 1040 は、HDD (Hard Disk Drive)、SSD (Solid State Drive)、メモリカード、又はROM (Read Only Memory) などを実現される補助記憶装置である。ストレージデバイス 1040 はモデル生成装置 10 の各機能（例えば訓練データ取得部 120、事前処理部 130、モデル生成部 140、及びモデル送信部 160）を実現するプログラムモジュールを記憶している。プロセッサ 1020 がこれら各プログラムモジュールをメモリ 1030 上に読み込んで実行することで、そのプログラムモジュールに対応する各機能が実現される。また、ストレージデバイス 1040 は訓練データ記憶部 110 及びモデル記憶部 150 としても機能する。
- [0055] 入出力インタフェース 1050 は、モデル生成装置 10 と各種入出力機器とを接続するためのインタフェースである。
- [0056] ネットワークインタフェース 1060 は、モデル生成装置 10 をネットワークに接続するためのインタフェースである。このネットワークは、例えばLAN (Local Area Network) やWAN (Wide Area Network) である。ネットワークインタフェース 1060 がネットワークに接続する方法は、無線接続であってもよいし、有線接続であってもよい。モデル生成装置 10 は、ネットワークインタフェース 1060 を介してデータ処理システム 20 と通信してもよい。
- [0057] なお、データ処理システム 20 のハードウェア構成も図 7 に示した例と同様である。そして、ストレージデバイスは、データ処理システム 20 の各機能（例えば記憶処理部 210、入力データ処理部 230、中間データ生成部 240、出力データ生成部 250、及び表示処理部 260）を実現するプログラムモジュールを記憶している。またストレージデバイスは、モデル記憶部 220 としても機能する。
- [0058] 図 8 は、モデル生成装置 10 が行うモデルの生成処理の第 1 例を示すフローチャートである。
- [0059] まずモデル生成装置 10 の訓練データ取得部 120 は、訓練データ記憶部 110 から訓練データを読み出す（ステップ S10）。この訓練データには

、第1訓練データ及び第2訓練データが含まれる。この状態において、第1訓練データ及び第2訓練データは変換ルールが適用されておらず、そのため、ダミーデータを加えられていない。

[0060] 次いでモデル生成部140は、第1訓練データを用いてモデルを訓練する（ステップS20）。以下、この訓練により生成されたモデルをベースモデルという。ベースモデルに対する入力データの値の数、及びベースモデルからの出力データの値の数は、それぞれ本来の数である。

[0061] 次いで、モデル生成部140は、ベースモデルを処理することにより、仮モデルを生成する（ステップS30）。この仮モデルは、入力データの数及び出力データの数それぞれが、変換ルールが適用された後の数となっている。

[0062] ここで、仮モデルの生成方法について、説明する。仮モデルは、例えば以下の方法によって生成される。なお、上記した第1訓練データの変換ルール、および中間データから出力データへの変換ルールは、この仮モデルの生成方法に従って定まる。

（1）ベースモデルのパラメータを示す行列において、特定の列、特定の行、もしくは任意の位置にダミーデータを追加する（図4～図6に示した例に対応）。

（2）ベースモデルのパラメータを示す行列において、特定の列、特定の行、もしくは特定の成分を相互に入れ替える（図11参照）。

（3）（1）及び（2）を組み合わせる。

[0063] ここで、（1）又は（3）において、仮モデルの訓練を行うことにより、仮モデルにおけるダミーデータの値が定まる。

[0064] 次いで事前処理部130は、第1訓練データおよび第2訓練データを変換ルールに従って変換する（ステップS40）。そしてモデル生成部140は、変換後の第1訓練データおよび第2訓練データを用いて仮モデルを訓練する（ステップS50）。その後、訓練後の仮モデルが基準を満たすか否かを判断する（ステップS60）。この基準は、例えば以下の2つの双方を含ん

でいる。

- [0065] 第1の基準は、訓練に用いられなかった実績データを訓練データ記憶部110から選択し、選択した実績データの入力データを仮モデルに入力し、その結果としての出力データと、選択した実績データのうち出力データに相当するデータと、の差が基準以下になることである。
- [0066] 第2の基準は、訓練に用いられなかった異常なデータ（第2訓練データに相当）の入力データを仮モデルに入力し、その結果として得られた出力データを構成するすべての値が、本来の出力データが取り得る値の範囲内にあることである。
- [0067] 訓練後の仮モデルが基準を満たさなかった場合（ステップS60：No）、ステップS20に戻る。すなわちモデル生成装置10は、仮モデルが基準を満たすまで、ステップS20～ステップS50に示した処理を繰り返す。
- [0068] 訓練後の仮モデルが基準を満たした場合（ステップS60：Yes）、モデル生成部140は、仮モデルを正式なモデルとしてモデル記憶部150に記憶させる（ステップS70）。
- [0069] 図9は、モデル生成装置10が行うモデルの生成処理の第2例を示すフローチャートである。本図に示す例は、第1訓練データを変換ルールで変換した後（ステップS22）にモデルを訓練することにより、仮モデルを生成する（ステップS32）点を除いて、図8に示した例と同様である。
- [0070] 図10は、データ処理システム20が行う処理の一例を示すフローチャートである。まず入力データ処理部230は、入力データを取得する。例えば入力データ処理部230は、蓄電池30から入力データを取得し（ステップS110）、この入力データを変換ルールに従って変換する（ステップS120）。次いで中間データ生成部240は、変換後の入力データを、モデル記憶部220が記憶しているモデルに入力することにより、中間データを得る（ステップS130）。次いで出力データ生成部250は、中間データに所定の演算を行った後に所定の位置にある値を選択することにより、出力データを生成する（ステップS140）。その後、出力データ生成部250は

、表示処理部 260 に出力データを出力する（ステップ S150）。表示処理部 260 は、ディスプレイ 270 に出力データを表示させる。

[0071] 以上、本実施形態によれば、データ処理システム 20 が用いるモデルの出力データは、データ処理システム 20 の出力データとは異なる。このため、第三者がデータ処理システム 20 の入力データと出力データの組み合わせを複数得たとしても、これらの組み合わせから 20 が用いるモデルを推定しにくい。したがって、このモデルの秘匿性は高い。

[0072] 以上、図面を参照して本発明の実施形態について述べたが、これらは本発明の例示であり、上記以外の様々な構成を採用することもできる。

[0073] また、上述の説明で用いた複数のフローチャートでは、複数の工程（処理）が順番に記載されているが、各実施形態で実行される工程の実行順序は、その記載の順番に制限されない。各実施形態では、図示される工程の順番を内容的に支障のない範囲で変更することができる。また、上述の各実施形態は、内容が相反しない範囲で組み合わせることができる。

[0074] この出願は、2020年10月9日に提出された日本出願特願2020-171188号を基礎とする優先権を主張し、その開示の全てをここに取り込む。

符号の説明

[0075]	10	モデル生成装置
	20	データ処理システム
	30	蓄電池
	40	機器
	110	訓練データ記憶部
	120	訓練データ取得部
	130	事前処理部
	140	モデル生成部
	150	モデル記憶部
	160	モデル送信部

2 1 0	記憶処理部
2 2 0	モデル記憶部
2 3 0	入力データ処理部
2 4 0	中間データ生成部
2 5 0	出力データ生成部
2 6 0	表示処理部
2 7 0	ディスプレイ

請求の範囲

- [請求項1] 複数の値からなる入力データを取得し、当該入力データを変換ルールに従って変換する入力データ処理部と、
- 変換後の前記入力データを処理することにより、複数の値からなる複数行及び／または複数列のデータからなる中間データを出力する中間データ生成部と、
- 前記中間データに所定の演算を行った後に予め定められた位置にある少なくとも一つの値を選択することにより出力データを生成する出力データ生成部と、
- を備え、
- 前記中間データ生成部と、前記出力データ生成部とは異なる装置に設けられているデータ処理システム。
- [請求項2] 請求項1に記載のデータ処理システムにおいて、
- 前記入力データに含まれる複数の値は、対象物の状態を互いに異なる指標で測定した結果であり、
- 前記中間データ生成部は、訓練データを用いて生成されたモデルを用いて前記中間データを生成し、
- 前記変換ルールは、前記入力データにダミーデータを追加することにより、前記入力データに含まれる値の数を増やすことであるデータ処理システム。
- [請求項3] 請求項1又は2に記載のデータ処理システムにおいて、
- 異常な前記入力データが入力されたとき、前記中間データに含まれる複数の値のうち前記出力データとして選択されなかった値は、正常な前記入力データが入力されたときの前記出力データに含まれる値が取り得る範囲内にある、データ処理システム。
- [請求項4] 請求項3に記載のデータ処理システムにおいて、
- 前記中間データ生成部は、訓練データを用いて生成されたモデルを用いて前記中間データを生成し、

前記モデルを生成するときに用いられる訓練データは、正常なデータからなる第1訓練データ、及び一部に異常なデータを含む第2訓練データを含んでおり、

前記モデルは、

前記第1訓練データを用いて訓練され、その後に、前記変換ルールによって変換された後の前記第2訓練データを用いて訓練される工程を、複数回繰り返すことにより生成されている、データ処理システム。

[請求項5] 請求項4に記載のデータ処理システムにおいて、

前記第1訓練データは前記変換ルールによって変換されずに用いられるデータ処理システム。

[請求項6] 請求項2～5のいずれか一項に記載のデータ処理システムにおいて、

、

前記対象物は二次電池であるデータ処理システム。

[請求項7] 請求項1～6のいずれか一項に記載のデータ処理システムにおいて、

、

前記入力データ処理部は、前記入力データに所定の条件を満たす値が含まれていたときに、当該値を他の値に置換するデータ処理システム。

[請求項8] 請求項1～7のいずれか一項に記載のデータ処理システムが有する中間データ生成部で用いられるモデルを生成するモデル生成装置であって、

正常なデータからなる第1訓練データ、及び一部に異常なデータを含む第2訓練データを取得する訓練データ取得部と、

前記モデルを、前記第1訓練データを用いて訓練し、その後に、前記変換ルールによって変換された後の前記第2訓練データを用いて訓練する工程を、複数回繰り返すモデル生成部と、
を備えるモデル生成装置。

- [請求項9] コンピュータが、
- 複数の値からなる入力データを取得し、当該入力データを変換ルールに従って変換する入力データ処理と、
- 変換後の前記入力データを処理することにより、複数の値からなる複数行及び／または複数列のデータからなる中間データを出力する中間データ生成処理と、
- 前記中間データに所定の演算を行った後に予め定められた位置にある少なくとも一つの値を選択することにより出力データを生成する出力データ生成処理と、
- を行い、
- 前記中間データ生成処理と、前記出力データ生成処理とが、異なるコンピュータによって行われるデータ処理方法。
- [請求項10] 請求項1～7のいずれか一項に記載のデータ処理システムが有する中間データ生成部で用いられるモデルを生成するモデル生成方法であって、
- コンピュータに、
- 正常なデータからなる第1訓練データ、及び一部に異常なデータを含む第2訓練データを取得する訓練データ取得処理と、
- 前記モデルを、前記第1訓練データを用いて訓練し、その後に、前記変換ルールによって変換された後の前記第2訓練データを用いて訓練する工程を、複数回繰り返すモデル生成処理と、
- を行わせるモデル生成方法。
- [請求項11] コンピュータに、
- 複数の値からなる入力データを取得し、当該入力データを変換ルールに従って変換する入力データ処理機能と、
- 変換後の前記入力データを処理することにより、複数の値からなる複数行及び／または複数列のデータからなる中間データを出力する中間データ生成機能と、

を持たせるプログラム。

[請求項12]

コンピュータに、

請求項11に記載の中間データ生成機能によって生成された前記中間データに所定の演算を行った後に、予め定められた位置にある少なくとも一つの値を選択することにより出力データを生成する出力データ生成機能と、

を持たせるプログラム。

[請求項13]

コンピュータに、請求項1～7のいずれか一項に記載のデータ処理システムが有する中間データ生成部で用いられるモデルを生成する機能を持たせるプログラムであって、

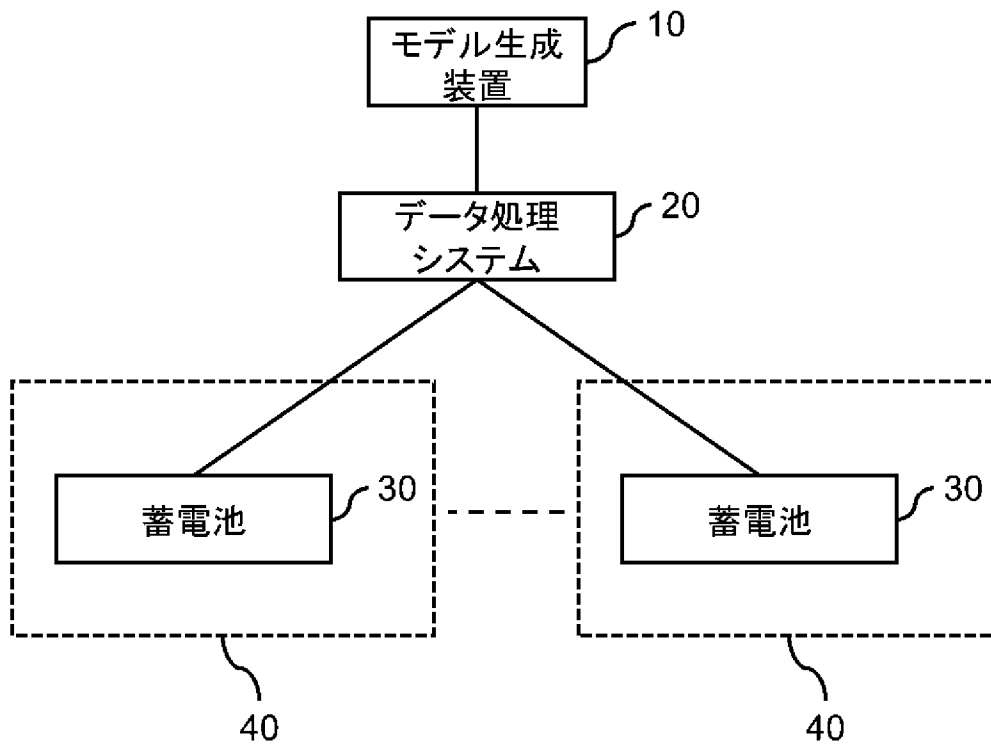
前記コンピュータに、

正常なデータからなる第1訓練データ、及び一部に異常なデータを含む第2訓練データを取得する訓練データ取得機能と、

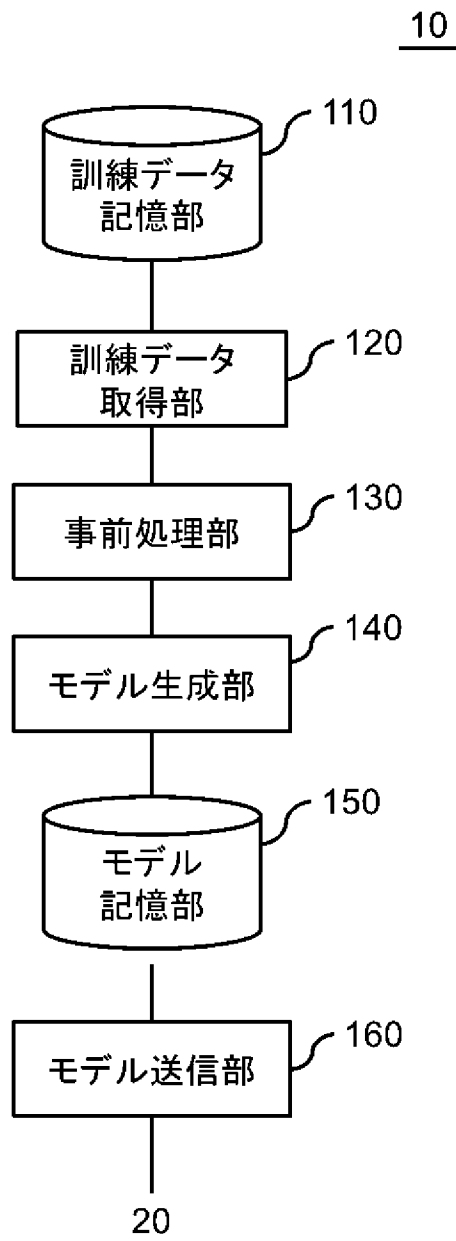
前記モデルを、前記第1訓練データを用いて訓練し、その後に、前記変換ルールによって変換された後の前記第2訓練データを用いて訓練する工程を、複数回繰り返すモデル生成機能と、

を持たせるプログラム。

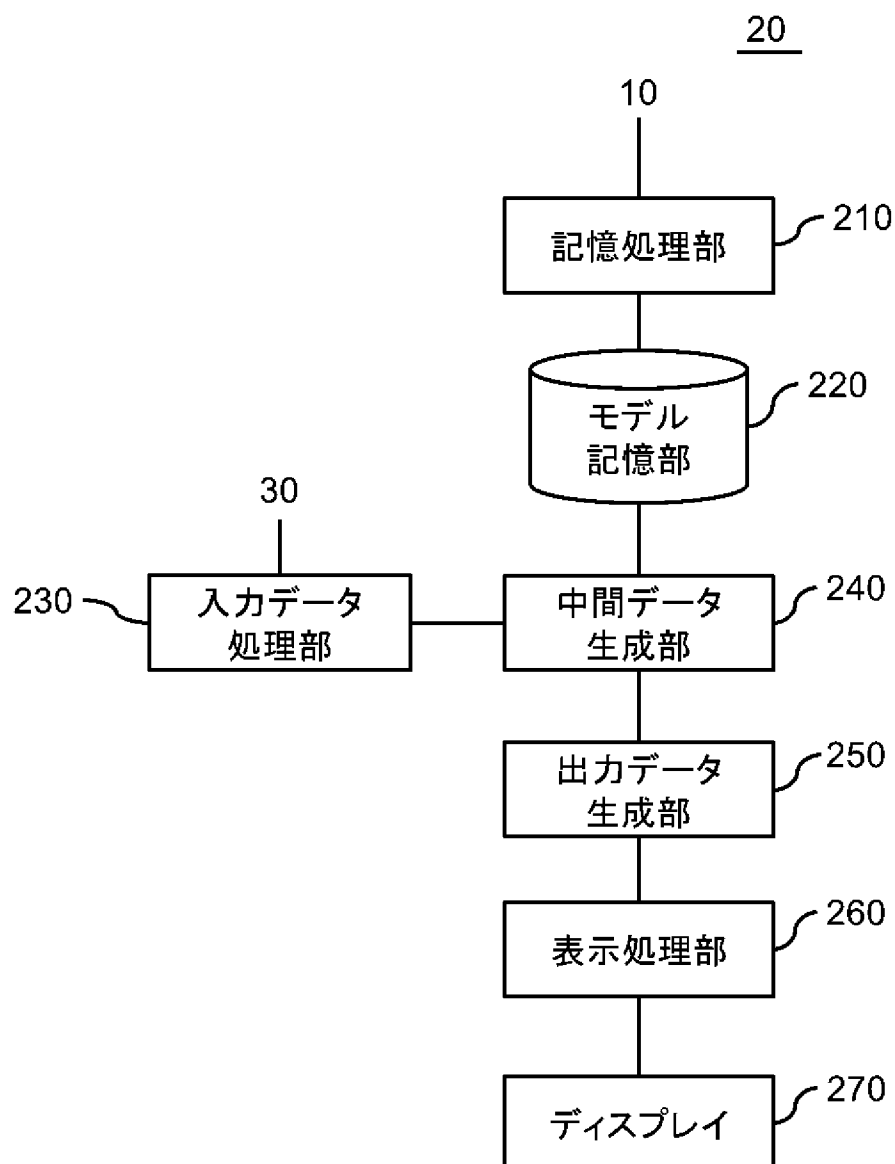
[図1]



[図2]



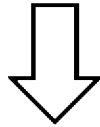
[図3]



[図4]

(A)

変換前の入力データ [1 1 1]



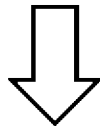
変換後の入力データ [1 0 1 1]



所定の値を付加

(B)

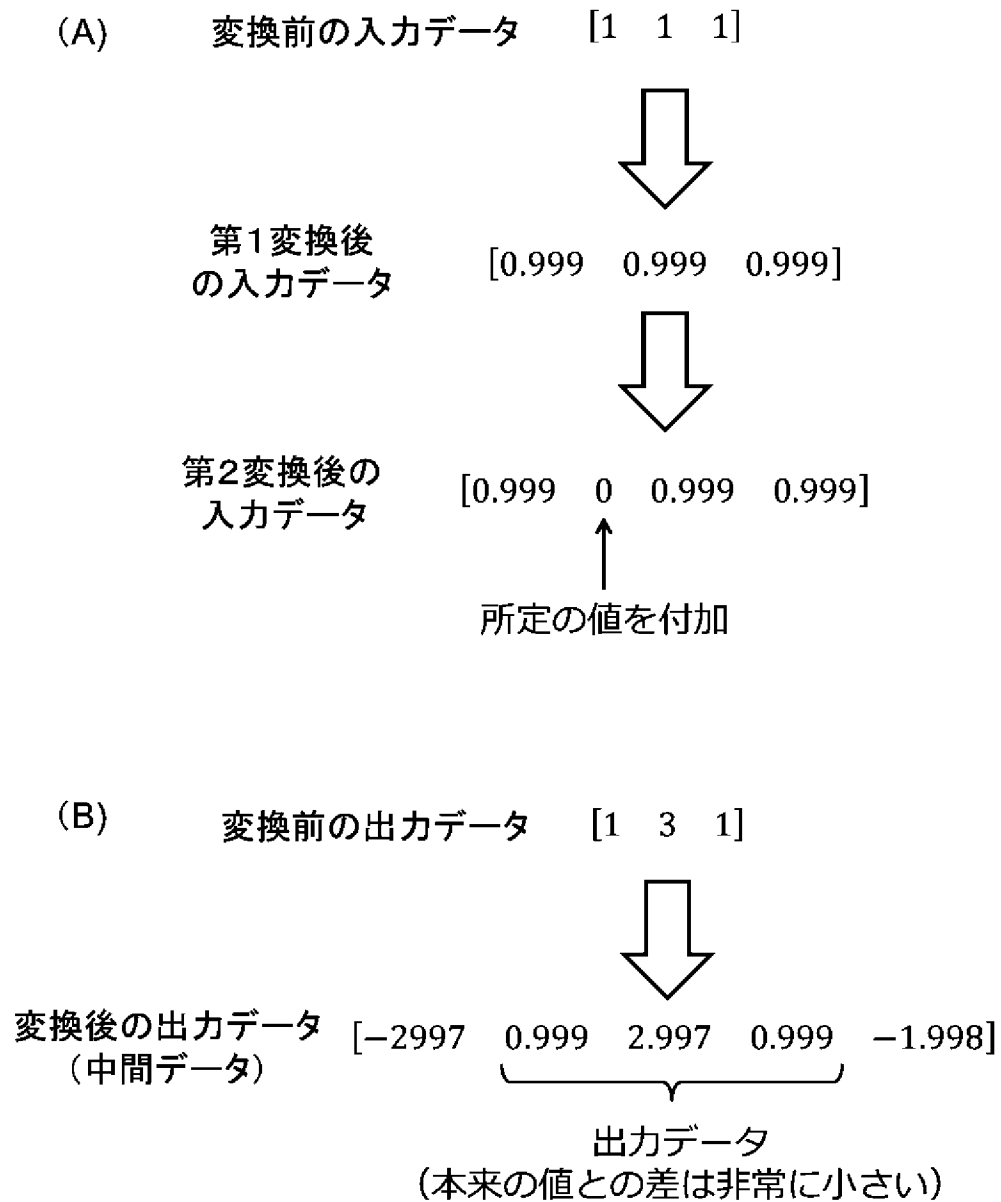
変換前の出力データ [1 3 1]



変換後の出力データ (中間データ) [-3 1 3 1 -2]

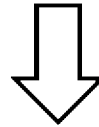
本来の出力データ

[図5]

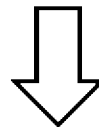


[図6]

(A) 変換前の入力データ [1 1 1]



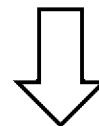
第1変換後の
入力データ [0.999 0.999 0.999]



第2変換後の
入力データ [0.999 0 0.999 0.999]

↑
所定の値を付加

(B) 変換前の出力データ [1 3 1]

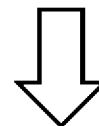


* バイアス分小さくなる
ようにモデルを訓練
(例: 第1訓練データを処理)

[-2997 0.999 1.997 0.999 -1.998]

+

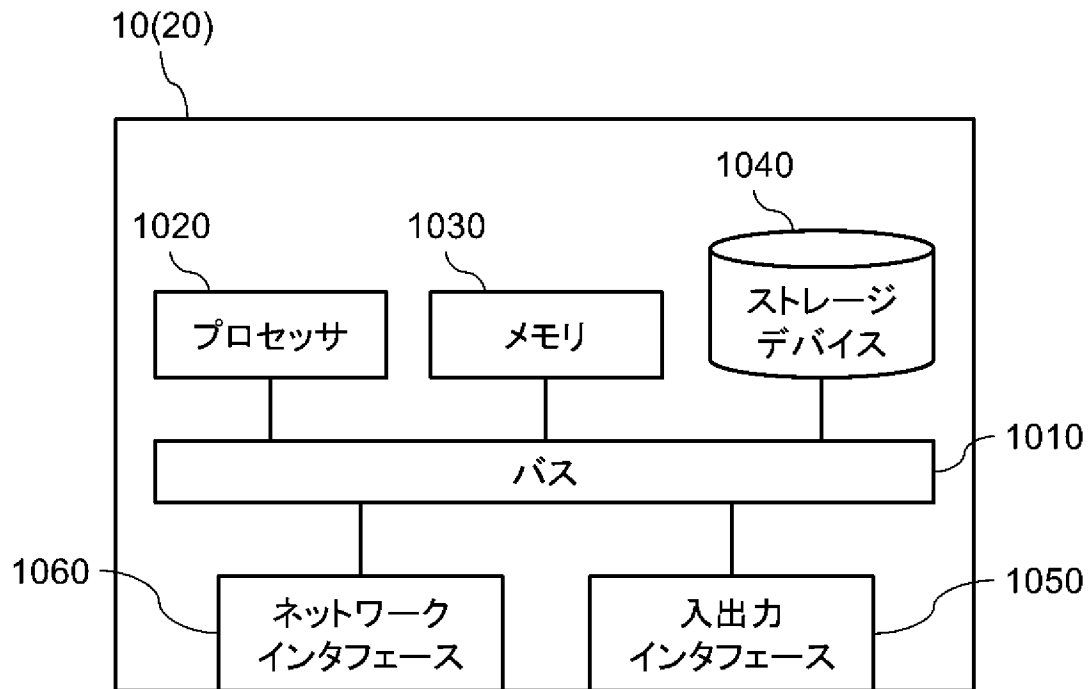
所定の演算 [0 0 1 0 0]



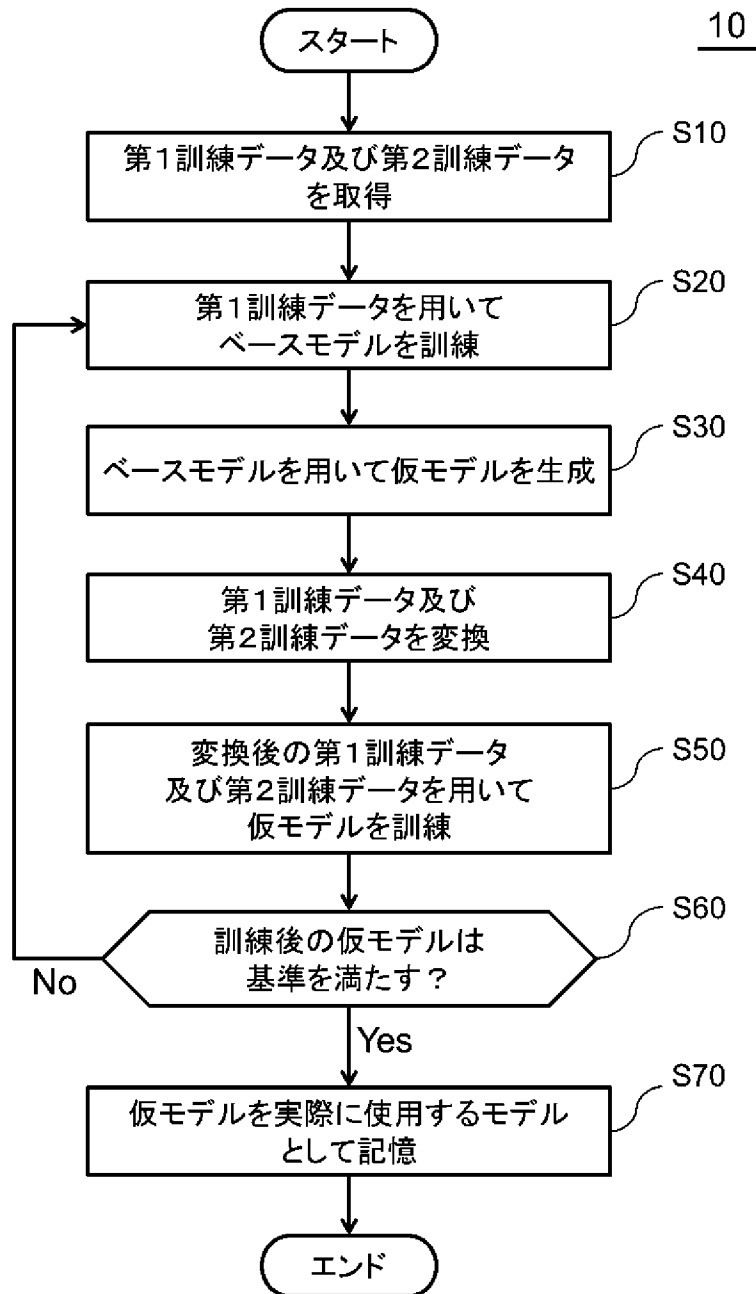
変換後の出力データ
(中間データ) [-2997 0.999 2.997 0.999 -1.998]

意味のある値
(本来の値との差は非常に小さい)

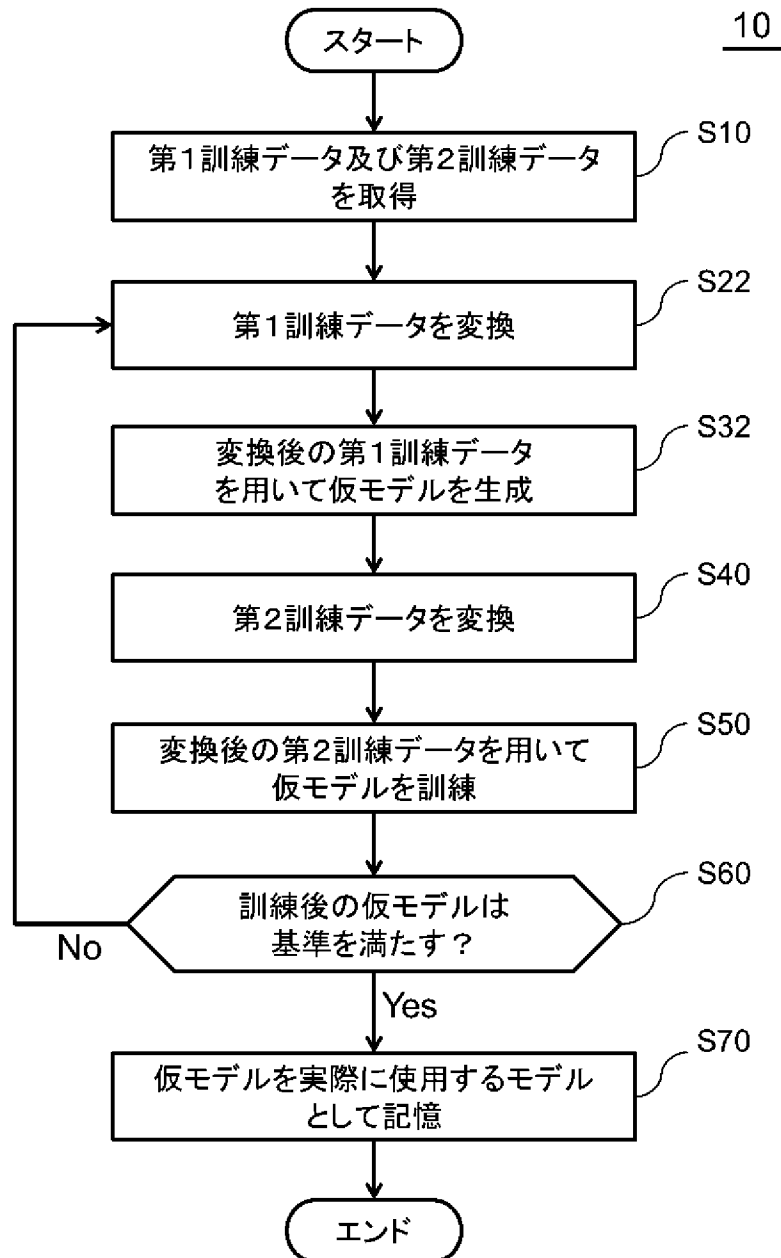
[図7]



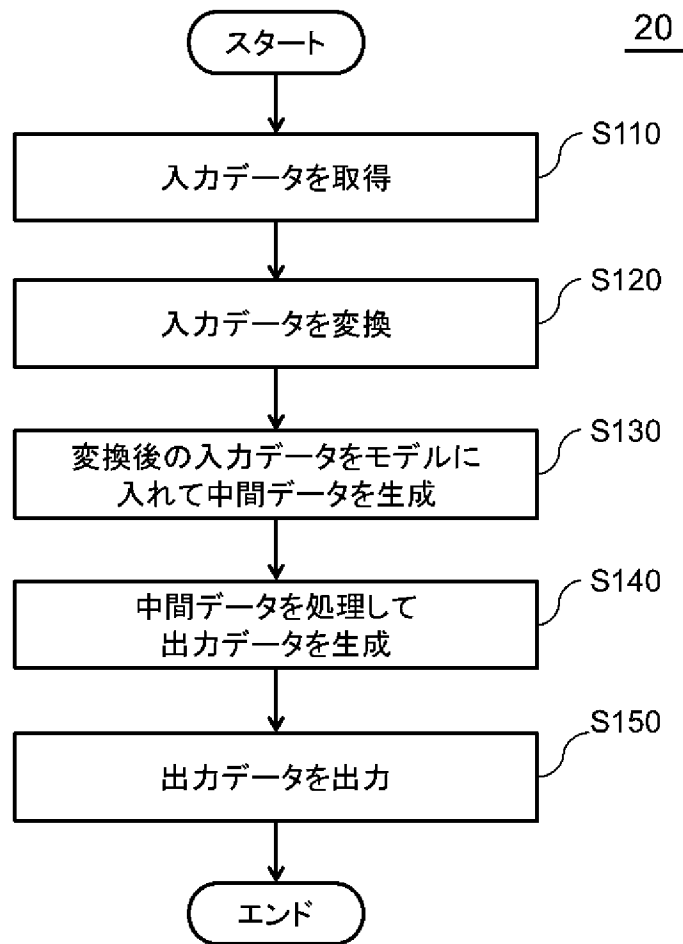
[図8]



[図9]



[図10]



[図11]

(A) ベースモデル

3	4	5
1	2	3
6	7	8

(B) 仮モデル

3	2	5
1	4	3
6	7	8

INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2021/036407

A. CLASSIFICATION OF SUBJECT MATTER**G06N 20/00**(2019.01)i

FI: G06N20/00

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)

G06N3/00-99/00

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 1971-2021
 Registered utility model specifications of Japan 1996-2021
 Published registered utility model applications of Japan 1994-2021

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

IEEE Xplore

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 2019/0044918 A1 (INTEL CORP.) 07 February 2019 (2019-02-07) paragraphs [0028], [0029], [0031], [0032], [0036]-[0038], fig. 2	1, 2, 9, 11, 12
Y		3, 4, 6-8, 10, 13
A		5
Y	JP 2017-207947 A (DNA CO., LTD.) 24 November 2017 (2017-11-24) paragraph [0036], fig. 4, 5	3, 4, 6-8, 10, 13
Y	GU, Zhongshu et al. Reaching Data Confidentiality and Model Accountability on the CalTrain, In: 2019 49th Annual IEEE/IFIP International Conference on Dependable Systems and Networks (DSN), 27 June 2019, IEEE, pp. 336-348, ISBN: 978-1-7281-0057-9, <DOI: 10.1109/DSN.2019.00044>.	4, 6-8, 10, 13
A	in particular, page 341, left column, lines 20-24	5
Y	US 2020/0164763 A1 (QUANTUMSCAPE CORP.) 28 May 2020 (2020-05-28) paragraphs [0133], [0137], [0139], claims 1, 8.	6-8, 10, 13

☒ 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

26 November 2021

Date of mailing of the international search report

07 December 2021

Name and mailing address of the ISA/JP

Japan Patent Office (ISA/JP)
 3-4-3 Kasumigaseki, Chiyoda-ku, Tokyo 100-8915
 Japan

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2021/036407**C. DOCUMENTS CONSIDERED TO BE RELEVANT**

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	BLOSSER, Gary et al. MAINTAINING K-ANONYMITY ON REAL-TIME DATA, In: Proceedings of the Sixth International Conference on Machine Learning and Cybernetics, 22 August 2007, IEEE, pp. 3012-3015, ISBN: 978-1-4244-0972-3, <DOI: 10.1109/ICMLC.2007.4370664> in particular, final six lines of section 1.1; first paragraph of section 3; tables 1, 2.	7, 8, 10, 13

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/JP2021/036407

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)	Publication date (day/month/year)
US	2019/0044918	A1	07 February 2019	CN 110322012 A paragraphs [0028], [0029], [0031], [0032], [0036]-[0038], fig. 2 EP 3547183 A1	
JP	2017-207947	A	24 November 2017	US 2017/0337449 A1 paragraph [0042], fig. 4, 5	
US	2020/0164763	A1	28 May 2020	WO 2019/017991 A1 paragraphs [0121], [0125], [0127]; claims 1, 8.	

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		
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）） G06N3/00-99/00		
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 1971-2021年 日本国実用新案登録公報 1996-2021年 日本国登録実用新案公報 1994-2021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） IEEE Xplore		
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	US 2019/0044918 A1 (INTEL CORPORATION) 07.02.2019 (2019-02-07) Paragraphs [0028], [0029], [0031], [0032], [0036]-[0038]; Figure 2.	1, 2, 9, 11, 12
Y		3, 4, 6-8, 10, 13
A		5
Y	JP 2017-207947 A (株式会社ディー・エヌ・エー) 24.11.2017 (2017-11-24) 段落[0036]; 図4, 5.	3, 4, 6-8, 10, 13
Y	GU, Zhongshu, et al., Reaching Data Confidentiality and Model Accountability on the CalTrain, In: 2019 49th Annual IEEE/IFIP International Conference on Dependable Systems and Networks (DSN), 2019.06.27, IEEE, Pages 336-348, ISBN: 978-1-7281-0057-9, <DOI: 10.1109/DSN.2019.00044>. Particularly, Page 341, left-hand column, lines 20-24.	4, 6-8, 10, 13
A		5
Y	US 2020/0164763 A1 (QUANTUMSCAPE CORPORATION) 28.05.2020 (2020-05-28) Paragraphs [0133], [0137], [0139]; Claims 1 and 8.	6-8, 10, 13
☒ 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 26. 11. 2021		国際調査報告の発送日 07. 12. 2021
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） 久保 光宏 5C 9189 電話番号 03-3581-1101 内線 3539

C. 関連すると認められる文献

引用文献の カテゴリ*	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
Y	BLOSSER, Gary, et al., MAINTAINING K-ANONYMITY ON REAL-TIME DATA, In: Proceedings of the Sixth International Conference on Machine Learning and Cybernetics, 2007.08.22, IEEE, Pages 3012-3015, ISBN: 978-1-4244-0972-3, <DOI: 10.1109/ICMLC.2007.4370664>. Particularly, Final six lines of section 1.1; First paragraph of section 3; Tables 1 and 2.	7, 8, 10, 13

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/JP2021/036407

引用文献			公表日	パテントファミリー文献	公表日
US	2019/0044918	A1	07.02.2019	CN 110322012 A Paragrahs [0028], [0029], [0031], [0032], [0036]- [0038]; Figure 2. EP 3547183 A1	
JP	2017-207947	A	24.11.2017	US 2017/0337449 A1 Paragraph [0042]; Figures 4 and 5.	
US	2020/0164763	A1	28.05.2020	WO 2019/017991 A1 Paragraphs [00121], [00125], [00127]; Claims 1 and 8.	