

(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
2021年6月24日(24.06.2021)



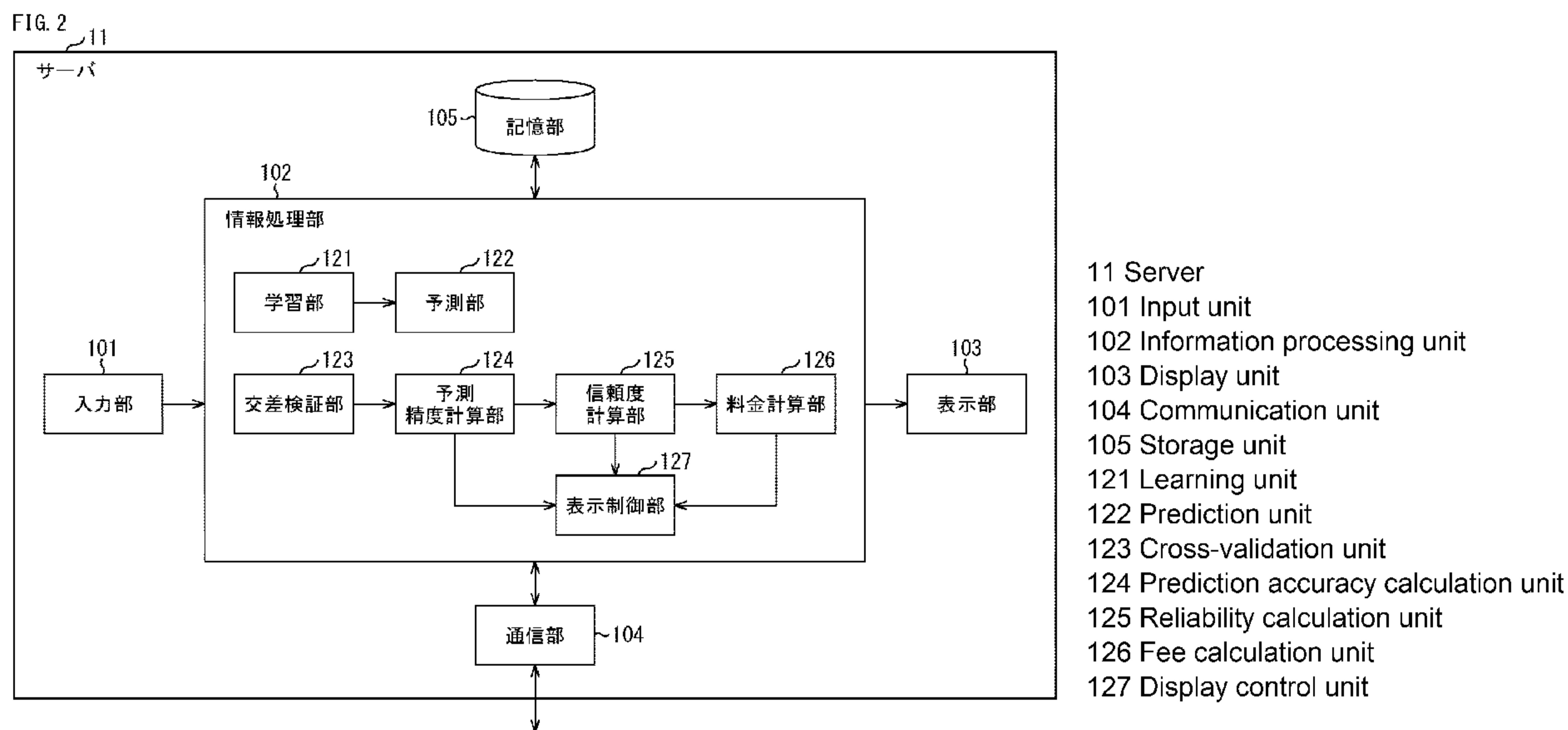
(10) 国際公開番号

WO 2021/124904 A1

- (51) 国際特許分類:
G06Q 10/04 (2012.01)
- (21) 国際出願番号: PCT/JP2020/044987
- (22) 国際出願日: 2020年12月3日(03.12.2020)
- (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 2019-227552 2019年12月17日(17.12.2019) JP
- (71) 出願人: ソニーグループ株式会社(SONY GROUP CORPORATION) [JP/JP]; 〒1080075 東京都港区港南1丁目7番1号 Tokyo (JP).
- (72) 発明者: 館野 啓(TATENO Kei); 〒1080075 東京都港区港南1丁目7番1号 ソニー株式会社内 Tokyo (JP).
- (74) 代理人: 西川 孝, 外(NISHIKAWA Takashi et al.); 〒1700013 東京都豊島区東池袋3丁目9番10号 池袋F Nビル4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, JP, 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,

(54) Title: INFORMATION PROCESSING DEVICE, INFORMATION PROCESSING METHOD, AND PROGRAM

(54) 発明の名称: 情報処理装置、情報処理方法、及び、プログラム



(57) Abstract: The present technology pertains to an information processing device, an information processing method, and a program with which it is possible to detect fraud in actual performance data that is used for calculating the prediction accuracy of a prediction model. The information processing device includes: a prediction accuracy calculation unit that calculates, on the basis of first actual performance data indicating actual performance values for a prediction target sample subject to prediction of a prediction model, a first prediction accuracy for the prediction target sample, and calculates, on the basis of second actual performance data indicating actual performance values for a non-prediction-target sample not subject to prediction of the prediction model, a second prediction accuracy for the non-prediction-target sample; and a reliability calculation unit that calculates, on the basis of the first prediction accuracy and the second prediction accuracy, reliability of the first actual performance data. The present technology can be applied, for example, to a server that provides

NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT,
QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL,
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))

prediction models.

(57) 要約 : 本技術は、予測モデルの予測精度の計算に用いる実績データの不正を検出することができるようにする情報処理装置、情報処理方法、及び、プログラムに関する。情報処理装置は、予測モデルの予測対象となる予測対象サンプルに対する実績値を示す第1の実績データに基づいて、前記予測対象サンプルに対する第1の予測精度を計算し、前記予測モデルの予測対象ではない予測対象外サンプルに対する実績値を示す第2の実績データに基づいて、前記予測対象外サンプルに対する第2の予測精度を計算する予測精度計算部と、前記第1の予測精度及び前記第2の予測精度に基づいて、前記第1の実績データの信頼度を計算する信頼度計算部とを備える。本技術は、例えば、予測モデルを提供するサーバに適用できる。

明 細 書

発明の名称： 情報処理装置、情報処理方法、及び、プログラム
技術分野

[0001] 本技術は、情報処理装置、情報処理方法、及び、プログラムに関し、特に、予測モデルの予測精度を計算する場合に用いて好適な情報処理装置、情報処理方法、及び、プログラムに関する。

背景技術

[0002] 従来、所定の事象を予測する各予測モデルの精度と、各予測モデルが提供された順序とに応じた貢献度に基づいて、各データサイエンティストへの報酬を計算する技術が提案されている（例えば、特許文献1参照）。

先行技術文献

特許文献

[0003] 特許文献1：特開2019-46456号公報

発明の概要

発明が解決しようとする課題

[0004] 例えば、予測モデルの利用料金を、特許文献1に記載の技術のように、予測精度に応じた成果報酬により設定する場合、予測モデルを提供する側（以下、モデル提供者と称する）が、予測モデルを利用する側（以下、モデル利用者と称する）から提供される実績値に基づいて、予測精度を計算することが想定される。しかしながら、この場合、モデル利用者が、料金を不当に下げようとして、実績値を示す実績データを改ざんしてモデル提供者に提供し、予測精度を下げようとするおそれがある。

[0005] 本技術は、このような状況に鑑みてなされたものであり、予測モデルの予測精度の計算に用いる実績データの不正を検出できるようにするものである。

課題を解決するための手段

[0006] 本技術の一側面の情報処理装置は、予測モデルの予測対象となる予測対象

サンプルに対する実績値を示す第1の実績データに基づいて、前記予測対象サンプルに対する第1の予測精度を計算し、前記予測モデルの予測対象ではない予測対象外サンプルに対する実績値を示す第2の実績データに基づいて、前記予測対象外サンプルに対する第2の予測精度を計算する予測精度計算部と、前記第1の予測精度及び前記第2の予測精度に基づいて、前記第1の実績データの信頼度を計算する信頼度計算部とを備える。

[0007] 本技術の一側面の情報処理方法は、情報処理装置が、予測モデルの予測対象となる予測対象サンプルに対する実績値を示す第1の実績データに基づいて、前記予測対象サンプルに対する第1の予測精度を計算し、前記予測モデルの予測対象ではない予測対象外サンプルに対する実績値を示す第2の実績データに基づいて、前記予測対象外サンプルに対する第2の予測精度を計算し、前記第1の予測精度及び前記第2の予測精度に基づいて、前記第1の実績データの信頼度を計算する。

[0008] 本技術の一側面のプログラムは、予測モデルの予測対象となる予測対象サンプルに対する実績値を示す第1の実績データに基づいて、前記予測対象サンプルに対する第1の予測精度を計算し、前記予測モデルの予測対象ではない予測対象外サンプルに対する実績値を示す第2の実績データに基づいて、前記予測対象外サンプルに対する第2の予測精度を計算し、前記第1の予測精度及び前記第2の予測精度に基づいて、前記第1の実績データの信頼度を計算する処理をコンピュータに実行させる。

[0009] 本技術の一側面によれば、予測モデルの予測対象となる予測対象サンプルに対する実績値を示す第1の実績データに基づいて、前記予測対象サンプルに対する第1の予測精度が計算され、前記予測モデルの予測対象ではない予測対象外サンプルに対する実績値を示す第2の実績データに基づいて、前記予測対象外サンプルに対する第2の予測精度が計算され、前記第1の予測精度及び前記第2の予測精度に基づいて、前記第1の実績データの信頼度が計算される。

図面の簡単な説明

[0010] [図1]本技術を適用した情報処理システムの一実施の形態を示すブロック図である。

[図2]サーバの構成例を示すブロック図である。

[図3]クライアントの構成例を示すブロック図である。

[図4]サーバにより実行される予測サービス提供処理を説明するためのフローチャートである。

[図5]学習用データの例を示す図である。

[図6]予測対象サンプルデータの例を示す図である。

[図7]予測結果データの例を示す図である。

[図8]実績データの例を示す図である。

[図9]予測対象外サンプルデータの例を示す図である。

[図10]交差検証用データの例を示す図である。

[図11]交差検証を説明するための図である。

[図12]モデル提供者用の表示画面の例を示す図である。

[図13]モデル利用者用の表示画面の例を示す図である。

[図14]コンピュータの構成例を示す図である。

発明を実施するための形態

[0011] 以下、本技術を実施するための形態について説明する。説明は以下の順序で行う。

1. 実施の形態
2. 変形例
3. その他

[0012] <<1. 実施の形態>>

図1乃至図13を参照して、本技術の実施の形態について説明する。

[0013] <情報処理システム1の構成例>

図1は、本技術を適用した情報処理システム1の一実施の形態を示している。

[0014] 情報処理システム1は、サーバ11、及び、クライアント12-1乃至ク

クライアント 1 2 - n を備える。サーバ 1 1、及び、クライアント 1 2 - 1 乃至クライアント 1 2 - n は、ネットワーク 1 3 を介して相互に接続され、相互に通信を行う。

[0015] なお、以下、クライアント 1 2 - 1 乃至クライアント 1 2 - n を個々に区別する必要がない場合、単にクライアント 1 2 と称する。

[0016] サーバ 1 1 は、予測モデルを用いて、各種の事象の予測を行うサービス（以下、予測サービスと称する）を提供するモデル提供者により使用される。なお、予測サービスにおいて予測対象となる事象は、特に限定されない。

[0017] 各クライアント 1 2 は、予測サービスを利用するモデル利用者により使用される。クライアント 1 2 は、例えば、パーソナルコンピュータ（PC）、スマートフォン、タブレット端末等の情報処理端末により構成される。

[0018] そして、サーバ 1 1 は、各クライアント 1 2 に対して、予測サービスを提供する。

[0019] 例えば、クライアント 1 2 は、所定の事象の予測を行う予測モデルの学習に用いる学習用データを、ネットワーク 1 3 を介してサーバ 1 1 に送信する。

[0020] サーバ 1 1 は、学習用データに基づいて、予測モデルを生成する。例えば、サーバ 1 1 は、ユーザの特徴量に基づいてユーザの行動を予測する予測モデルを生成する。

[0021] また、クライアント 1 2 は、所定の事象の発生確率等の予測対象となるサンプル（以下、予測対象サンプルと称する）に関するデータ（以下、予測対象サンプルデータと称する）を、ネットワーク 1 3 を介してサーバ 1 1 に送信する。なお、予測対象となるサンプルは、特に限定されず、例えば、人、物、空間、時間等が想定される。また、予測対象サンプルデータは、例えば、各予測対象サンプルの特徴量等を含む。

[0022] サーバ 1 1 は、予測モデル、及び、予測対象サンプルデータに基づいて、各予測対象サンプルの所定の事象の発生確率等を予測する。サーバ 1 1 は、予測結果を示すデータ（以下、予測結果データと称する）を、ネットワーク

13を介してクライアント12に送信する。

[0023] モデル利用者は、予測結果データを利用して、各種の行動を行う。例えば、モデル利用者は、予測対象サンプルである各ユーザの行動予測に基づいて、各ユーザに適したサービスやプロモーション等を提供する。

[0024] そして、クライアント12は、予測対象サンプルに対する所定の事象に関する実際の結果（以下、実績値と称する）を示す実績データ（以下、予測対象実績データと称する）を、ネットワーク13を介してサーバ11に送信する。また、クライアント12は、予測対象ではないサンプル（以下、予測対象外サンプルと称する）に関するデータ（以下、予測対象外サンプルデータ）、及び、予測対象外サンプルに対する実績データ（以下、予測対象外実績データと称する）を、ネットワーク13を介してサーバ11に送信する。予測対象外サンプルデータは、例えば、各予測対象外サンプルの特徴量等を含む。

[0025] サーバ11は、予測対象サンプルデータ、予測対象実績データ、予測対象外サンプルデータ、及び、予測対象外実績データに基づいて、予測対象サンプル及び予測対象外サンプルに対する予測精度を計算する。また、サーバ11は、予測対象サンプル及び予測対象外サンプルに対する予測精度に基づいて、実績データの信頼度を計算する。さらに、サーバ11は、予測対象サンプル及び予測対象外サンプルに対する予測精度、並びに、実績データの信頼度に基づいて、モデル利用者に課金する料金を設定する。サーバ11は、料金及び料金の根拠を示す画面を表示させるためのデータ（以下、表示制御データと称する）を、ネットワーク13を介してクライアント12に送信する。

[0026] クライアント12は、表示制御データに基づいて、料金及び料金の根拠を示す画面を表示する。そして、モデル利用者は、例えば、クライアント12を用いて、料金の支払い手続や問合せ等を行う。

[0027] <サーバ11の構成例>

図2は、サーバ11の構成例を示すブロック図である。

- [0028] サーバ11は、入力部101、情報処理部102、表示部103、通信部104、及び、記憶部105を備える。
- [0029] 入力部101は、キーボード、マウス、マイクロフォン等の入力デバイスを備え、入力されたデータを情報処理部102に供給する。
- [0030] 情報処理部102は、学習部121、予測部122、交差検証部123、予測精度計算部124、信頼度計算部125、料金計算部126、及び、表示制御部127を備える。
- [0031] 学習部121は、クライアント12から取得した学習用データに基づいて、予測モデルの学習を行う。学習部121は、生成した予測モデルを予測部122に供給したり、記憶部105に記憶させたりする。
- [0032] 予測部122は、予測モデル、及び、クライアント12から取得した予測対象サンプルデータに基づいて、予測対象サンプルの所定の事象に関する予測を行う。予測部122は、予測結果を示す予測結果データを生成し、通信部104を介してクライアント12に送信したり、記憶部105に記憶させたりする。
- [0033] 交差検証部123は、クライアント12から取得した予測対象サンプルデータ、予測対象実績データ、予測対象外サンプルデータ、及び、予測対象外実績データを用いて交差検証を行う。これにより、交差検証の対象となる予測モデルと同じ事象の予測を行う予測モデルである交差検証モデルが生成される。交差検証部123は、生成した交差検証モデルを予測精度計算部124に供給する。
- [0034] 予測精度計算部124は、クライアント12から取得した予測対象サンプルデータ、予測対象実績データ、予測対象外サンプルデータ、及び、予測対象外実績データ、並びに、交差検証部123から取得した交差検証モデルを用いて、各サンプルに対する予測精度を計算する。例えば、予測精度計算部124は、交差検証モデルを用いて、各サンプルに対する予測確率（以下、CV予測確率と称する）を計算する。また、予測精度計算部124は、予測対象サンプルに対するCV予測確率、及び、予測対象実績データに基づいて

、予測対象サンプルに対する予測精度を計算する。さらに、予測精度計算部 124 は、予測対象外サンプルに対する CV 予測確率、及び、予測対象外実績データに基づいて、予測対象外サンプルに対する予測精度を計算する。予測精度計算部 124 は、予測精度の計算結果を示すデータを信頼度計算部 125 及び表示制御部 127 に供給する。

[0035] 信頼度計算部 125 は、予測対象サンプルに対する予測精度と、予測対象外サンプルに対する予測精度とを比較することにより、クライアント 12 から取得した実績データに対する信頼度を計算する。また、予測精度計算部 124 は、信頼度に基づいて、実績データの不正の有無を判定する。信頼度計算部 125 は、予測精度の計算結果、信頼度の計算結果、及び、実績データの不正の有無の判定結果を示すデータを料金計算部 126 及び表示制御部 127 に供給する。

[0036] 料金計算部 126 は、予測対象サンプルに対する予測精度、予測対象外サンプルに対する予測精度、実績データの信頼度、及び、実績データの不正の有無の判定結果等に基づいて、モデル利用者に課金する料金を計算する。料金計算部 126 は、料金の計算結果を示すデータを表示制御部 127 に供給する。

[0037] 表示制御部 127 は、表示部 103 の表示処理の制御を行う。例えば、表示制御部 127 は、モデル利用者に課金する料金及び料金の根拠を示す画面の表示部 103 による表示を制御する。

[0038] また、表示制御部 127 は、クライアント 12 の表示処理の制御を行う。例えば、表示制御部 127 は、モデル利用者に課金する料金及び料金の根拠を示す画面を表示させるための表示制御データを生成し、通信部 104 を介してクライアント 12 に送信する。これにより、表示制御部 127 は、モデル利用者に課金する料金及び料金の根拠を示す画面をクライアント 12 に表示させる。

[0039] 表示部 103 は、ディスプレイ等の表示デバイスを備え、例えば、各種の情報や操作画面等の表示を行う。

- [0040] 通信部104は、各種の通信デバイスを備え、ネットワーク13を介して、各クライアント12と通信を行う。なお、通信部104の通信方式は、特に限定されず、通信部104が複数の通信方式に対応していてもよい。
- [0041] 記憶部105は、例えば、不揮発性及び揮発性のメモリを備え、サーバ11の処理に使用される各種のプログラムやデータを記憶する。
- [0042] <クライアント12の構成例>
図3は、クライアント12の構成例を示すブロック図である。
- [0043] クライアント12は、入力部201、情報処理部202、表示部203、通信部204、及び、記憶部205を備える。
- [0044] 入力部201は、キーボード、マウス、マイクロフォン等の入力デバイスを備え、入力されたデータを情報処理部202に供給する。
- [0045] 情報処理部202は、演算処理部221及び表示制御部222を備える。
- [0046] 演算処理部221は、各種の演算を行ったり、クライアント12の各部の制御を行ったりする。例えば、演算処理部221は、記憶部205に記憶されているログデータ等に基づいて、学習用データ、予測対象サンプルデータ、予測対象外サンプルデータ、予測対象実績データ、及び、予測対象外実績データ等を生成し、通信部204を介してサーバ11に送信する。
- [0047] 表示制御部222は、表示部203の表示処理の制御を行う。例えば、表示制御部222は、モデル利用者に課金する料金及び料金の根拠を示す画面の表示部203による表示を制御する。
- [0048] 表示部203は、ディスプレイ等の表示デバイスを備え、例えば、各種の情報や操作画面等の表示を行う。
- [0049] 通信部204は、各種の通信デバイスを備え、ネットワーク13を介して、サーバ11と通信を行う。なお、通信部204の通信方式は、特に限定されず、通信部204が複数の通信方式に対応していてもよい。
- [0050] 記憶部205は、例えば、不揮発性及び揮発性のメモリを備え、クライアント12の処理に使用される各種のプログラムやデータを記憶する。例えば、記憶部205は、クライアント12が提供するサービスを利用するユーザ

の個人情報や行動の履歴等を示すログデータを記憶する。

[0051] <予測サービス提供処理>

次に、図4のフローチャートを参照して、サーバ11により実行される予測サービス提供処理について説明する。

[0052] なお、以下、モデル利用者が提供するサービスを利用するユーザが、所定の期間内（例えば、1ヶ月以内）に商品を購入するか否かの予測を行うサービスをモデル提供者が提供する場合を具体例に挙げて説明する。

[0053] ステップS1において、学習部121は、モデル学習を行う。

[0054] 具体的には、例えば、モデル利用者は、クライアント12の入力部201を介して、学習用データの送信の指示を入力する。演算処理部221は、記憶部205に記憶されているログデータに基づいて、学習用データを生成する。

[0055] 図5は、学習用データの例を示している。学習用データは、学習対象となるユーザ（以下、学習対象ユーザと称する）に関するデータを含む。例えば、学習用データは、“ID”、ユーザの特徴量、及び、“購入実績”を含む。ユーザの特徴量は、例えば、“性別”、“年齢”、及び、“過去の購入回数”を含む。

[0056] “ID”は、モデル利用者が提供するサービスを利用するユーザを一意に識別するために、各ユーザに付与されるIDを示す。この例では、学習対象ユーザに1000番台のIDが付与されている。

[0057] “性別”は、ユーザの性別を示す。

[0058] “年齢”は、ユーザの年齢を示す。

[0059] “過去の購入回数”は、所定の購入回数カウント期間内にユーザが商品を購入した回数を示す。購入回数カウント期間は、例えば、所定の基準日より第1の期間前（例えば、1年1ヶ月前）から第2の期間前（例えば、1ヶ月前）の前日までの期間に設定される。基準日は、例えば、現在の日付に設定される。

[0060] “購入実績”は、所定の購入実績対象期間内にユーザが商品を購入したか

否かを示す。例えば、”購入実績”は、購入実績対象期間内にユーザが商品を購入した場合、購入回数に関わらず、1に設定され、購入実績対象期間内にユーザが商品を購入しなかった場合、0に設定される。購入実績対象期間は、例えば、基準日より上述した第2の期間前から基準日までの期間に設定される。

[0061] 演算処理部221は、生成した学習用データを、通信部204を介してサーバ11に送信する。

[0062] これに対して、サーバ11の情報処理部102は、ネットワーク13及び通信部104を介して、学習用データを受信し、記憶部105に記憶させる。

[0063] このサーバ11とクライアント12との間の学習用データの授受には、例えば、Web API (Application Programming Interface) 等が用いられる。

[0064] 学習部121は、記憶部105から学習用データを読み出し、学習用データを用いて機械学習を行う。これにより、例えば、ユーザが基準日以降の所定の期間内（例えば、1ヶ月以内）に商品を購入する確率を予測する予測モデルが生成される。

[0065] なお、予測モデルに用いるモデルは、特に限定されず、例えば、Deep Neural Networkやロジスティック回帰等の一般的な機械学習のモデルが用いられる。

[0066] 学習部121は、生成した予測モデルを記憶部105に記憶させる。

[0067] ステップS2において、予測部122は、予測モデルを用いて、予測を行う。

[0068] 具体的には、例えば、モデル利用者は、クライアント12の入力部201を介して、予測対象サンプルデータの生成及び送信の指示を入力する。演算処理部221は、記憶部205に記憶されているログデータに基づいて、予測対象サンプルデータを生成する。

[0069] 図6は、予測対象サンプルデータの例を示している。予測対象サンプルデ

ータは、予測対象となるサンプルであるユーザ（以下、予測対象ユーザと称する）に関するデータであり、図5の学習用データから”購入実績”を除いた項目を含む。なお、この例では、予測対象ユーザに2000番台のIDが付与されている。

[0070] 演算処理部221は、生成した予測対象サンプルデータを、通信部204を介してサーバ11に送信する。

[0071] これに対して、サーバ11の情報処理部102は、ネットワーク13及び通信部104を介して、予測対象サンプルデータを受信し、記憶部105に記憶させる。

[0072] このサーバ11とクライアント12との間の予測対象サンプルデータの授受には、例えば、Web API等が用いられる。

[0073] 予測部122は、予測モデル及び予測対象サンプルデータを記憶部105から読み出し、予測モデルに予測対象サンプルデータを入力する。予測モデルは、予測対象ユーザが所定の予測対象期間内に商品を購入する確率を予測し、予測結果を示す予測結果データを出力する。予測対象期間は、例えば、翌日以降の所定の期間内（例えば、1ヶ月以内）に設定される。

[0074] 図7は、予測結果データの例を示している。予測結果データは、”ID”及び”予測確率”の項目を含む。

[0075] ”予測確率”は、予測対象期間内にユーザが商品を購入する確率の予測値を示す。

[0076] ステップS3において、予測部122は、予測結果を提供する。具体的には、予測部122は、通信部104を介して、予測結果データをクライアント12に送信する。また、予測部122は、予測結果データを記憶部105に記憶させる。

[0077] これに対して、クライアント12の情報処理部202は、ネットワーク13及び通信部204を介して、予測結果データを受信し、記憶部205に記憶させる。

[0078] また、例えば、モデル利用者は、予測結果データに基づいて、ユーザに商

品の購入を促すための施策を実行する。例えば、モデル利用者は、予測確率の高いユーザに対して優先的に商品のプロモーションを行う等の施策を実行する。そして、モデル利用者は、入力部201を介して、各ユーザの購入実績をクライアント12に入力する。演算処理部221は、記憶部205に記憶されているログデータに、各ユーザの購入実績を記録する。

[0079] ステップS4において、サーバ11は、実績データを取得する。

[0080] 具体的には、例えば、モデル利用者は、クライアント12の入力部201を介して、実績データの生成及び送信の指示を入力する。演算処理部221は、記憶部205に記憶されているログデータに基づいて、実績データを生成する。

[0081] 図8は、実績データの例を示している。実績データは、“ID”及び“購入実績”を含む。

[0082] “購入実績”は、予測対象期間内に実際にユーザが商品を購入したか否かを示す。例えば、“購入実績”は、ユーザが予測対象期間内に商品を購入した場合、購入した回数に関わらず、1に設定され、ユーザが予測対象期間内に商品を購入しなかった場合、0に設定される。

[0083] なお、実績データには、予測対象ユーザの実績データ（予測対象実績データ）だけでなく、予測対象外ユーザの実績データ（予測対象外実績データ）が含まれる。予測対象外ユーザには、例えば、予測対象から外れたユーザや、モデル利用者が提供するサービスを新規に利用したユーザ等が含まれる。

[0084] なお、例えば、予測対象外ユーザに、学習対象ユーザを含めるようにしてもよい。

[0085] なお、図8の例では、予測対象ユーザに上述したように2000番台のIDが付与され、予測対象外ユーザに3000番台のIDが付与されている。

[0086] また、演算処理部221は、記憶部205に記憶されているログデータに基づいて、予測対象外ユーザに関するデータである予測対象外サンプルデータを生成する。予測対象外サンプルデータは、例えば、図9に示されるように、図6の予測対象サンプルデータと同様の項目を含む。

- [0087] 演算処理部 221 は、実績データ及び予測対象外サンプルデータを、通信部 204 を介してサーバ 11 に送信する。
- [0088] これに対して、サーバ 11 の情報処理部 102 は、ネットワーク 13 及び通信部 104 を介して、実績データ及び予測対象外サンプルデータを受信し、記憶部 105 に記憶させる。
- [0089] このサーバ 11 とクライアント 12 との間の実績データ及び予測対象外サンプルデータの授受には、例えば、Web API 等が用いられる。
- [0090] ステップ S5 において、交差検証部 123 は、実績データに基づいて、交差検証を行う。
- [0091] 具体的には、交差検証部 123 は、予測対象サンプルデータ、予測対象外サンプルデータ、及び、実績データを記憶部 105 から読み出す。交差検証部 123 は、予測対象サンプルデータ、予測対象外サンプルデータ、及び、実績データに基づいて、交差検証に用いる交差検証用データを生成する。
- [0092] 図 10 は、交差検証用データの例を示している。交差検証用データは、“ID”、“性別”、“年齢”、“過去の購入回数”、及び、“購入実績”を含む。すなわち、交差検証用データは、図 6 の予測対象サンプルデータと、図 8 の実績データのうち予測対象ユーザに対するデータ（予測対象実績データ）とを組み合わせたデータ（以下、予測対象サンプルデータセットと称する）を含む。また、交差検証用データは、図 9 の予測対象外サンプルデータと、図 8 の実績データのうち予測対象外ユーザに対するデータ（予測対象外実績データ）とを組み合わせたデータ（以下、予測対象外サンプルデータセットと称する）を含む。
- [0093] 交差検証部 123 は、交差検証用データに基づいて交差検証を行うことにより、予測モデルと同じ事象（例えば、ユーザが予測対象期間内に商品を購入する確率）の予測を行う予測モデルである交差検証モデルを生成する。
- [0094] 具体的には、例えば、交差検証部 123 は、交差検証用データに含まれるデータを、ID が奇数のユーザのデータと ID が偶数のユーザのデータの 2 グループに分割する。

- [0095] 次に、交差検証部 1 2 3 は、例えば、図 1 1 に示されるように、ID が奇数のユーザのグループ（以下、奇数グループと称する）のデータを訓練用データとし、ID が偶数のユーザのグループ（以下、偶数グループと称する）を検証用データとして、1 回目の交差検証を行う。なお、図 1 1 の CV 1 の欄が、1 回目の交差検証のデータの種別を示し、“Tr” は訓練用データを示し、“Te” は検証用データを示す。
- [0096] 次に、交差検証部 1 2 3 は、偶数グループのデータを訓練用データとし、奇数グループを検証用データとして、2 回目の交差検証を行う。なお、図 1 1 の CV 2 の欄が、2 回目の交差検証のデータの種別を示している。
- [0097] このように交差検証が行われることにより、交差検証モデルが生成される。
- [0098] 交差検証部 1 2 3 は、生成した交差検証モデルを予測精度計算部 1 2 4 に供給する。
- [0099] なお、交差検証用データの分割数を 3 以上に設定してもよい。この場合、分割後の各グループに、予測対象サンプルデータセット内のデータ及び予測対象外サンプルデータセット内のデータの両方が含まれるように、予測対象サンプルデータセット及び予測対象外サンプルデータセットが分散される。
- [0100] ステップ S 6 において、予測精度計算部 1 2 4 は、交差検証の結果に基づいて、予測精度を計算する。
- [0101] 具体的には、予測精度計算部 1 2 4 は、予測対象サンプルデータ、予測対象外サンプルデータ、及び、実績データを記憶部 1 0 5 から読み出す。そして、予測精度計算部 1 2 4 は、交差検証モデルに予測対象サンプルデータを入力することにより、予測対象ユーザが予測対象期間内に商品を購入する確率（以下、CV 予測確率と称する）を予測する。また、予測精度計算部 1 2 4 は、交差検証モデルに予測対象外サンプルデータを入力することにより、予測対象外ユーザが予測対象期間内に商品を購入する確率（CV 予測確率）を予測する。
- [0102] 図 1 1 には、各ユーザの CV 予測確率の例が示されている。

- [0103] 次に、予測精度計算部 124 は、予測対象ユーザの CV 予測確率、及び、購入実績に基づいて、予測対象ユーザに対する予測精度 α_p を計算する。また、予測精度計算部 124 は、予測対象外ユーザの CV 予測確率、及び、購入実績に基づいて、予測対象外ユーザに対する予測精度 α_q を計算する。
- [0104] なお、予測精度 α_p 及び予測精度 α_q の計算方法は、特に限定されない。例えば、AR (Accuracy Ratio) 値、Logistic Loss 等が用いられる。なお、例えば、損失関数が用いられる場合、その逆数が予測精度として用いられる。
- [0105] 予測精度計算部 124 は、予測精度 α_p 及び予測精度 α_q を示すデータを信頼度計算部 125 及び表示制御部 127 に供給する。
- [0106] ステップ S7 において、信頼度計算部 125 は、予測精度の計算結果に基づいて、実績データの信頼度を計算する。具体的には、信頼度計算部 125 は、予測精度 α_p と予測精度 α_q とを比較することより、実績データ、特に、予測対象ユーザの実績データの信頼度 R を計算する。例えば、信頼度計算部 125 は、次式 (1) により、信頼度 R を計算する。
- [0107]
$$R = \alpha_p / \alpha_q \quad \dots (1)$$
- [0108] ここで、予測モデルの予測対象ユーザに対する予測精度の真の値を α_f とした場合、実績データが改ざんされていない場合、予測精度 α_p は、予測精度 α_f とほぼ等しくなる。また、予測精度 α_p と予測精度 α_q の間には、ほとんど差が生じないと想定されるため、予測精度 α_q も予測精度 α_f とほぼ等しくなる。従って、信頼度 R は、1 に近い値になる。
- [0109] 一方、モデル利用者が、予測サービスの利用料金を不当に下げるために、予測対象ユーザに対する予測モデルの予測精度を下げようとした場合、例えば、予測対象ユーザの実績データと予測確率とを乖離させるように、予測対象ユーザの実績データを改ざんすることが想定される。換言すれば、モデル利用者が、予測対象ユーザの実績データと予測確率との差が大きくなるように、予測対象ユーザの実績データを改ざんすることが想定される。この場合、予測精度 α_p は、予測精度 α_f よりかなり低い値となる。
- [0110] 一方、例えば、予測対象外ユーザの実績データが予測モデルの再学習に用

いられるとした場合、モデル利用者が予測対象外ユーザの実績データも改ざんしてしまうと、予測モデルの再学習の精度が低下する。そのため、この場合、モデル利用者には、予測対象外ユーザの実績データを改ざんするインセンティブが働かない。従って、予測精度 α_q は、予測精度 α_f とほぼ同じ値のままとなると想定される。

[0111] 従って、モデル利用者が、予測サービスの利用料金を不当に下げようとした場合、予測精度 α_q に対して予測精度 α_p が小さくなり、信頼度 R が低下すると想定される。

[0112] ステップS8において、信頼度計算部125は、不正検出パラメータを設定する。不正検出パラメータを β とすると、不正検出パラメータ β は、 $0 < \beta < 1$ の範囲内で設定される。

[0113] 例えば、信頼度計算部125は、不正検出パラメータ β を所定のデフォルト値に設定する。或いは、例えば、信頼度計算部125は、不正検出パラメータ β を、前回図4のフローチャートに従って、予測サービス提供処理を行ったときの最終値に設定する。

[0114] ステップS9において、信頼度計算部125は、実績データに不正があるか否かを判定する。例えば、信頼度計算部125は、次式(2)により、不正検出パラメータ β に基づく可変の閾値を用いて、実績データに不正があるか否かを判定する。

[0115] $R < 1 - \beta$ (ただし、 $0 < \beta < 1$) . . . (2)

[0116] 例えば、信頼度計算部125は、式(2)が満たされる場合、実績データに不正があると判定し、式(2)が満たされない場合、実績データに不正がないと判定する。そして、実績データに不正がないと判定された場合、処理はステップS10に進む。

[0117] ステップS10において、サーバ11は、予測精度に基づく成果報酬により料金を計算する。

[0118] 具体的には、信頼度計算部125は、予測精度 α_p 、予測精度 α_q 、及び、実績データの不正の有無の判定結果を示すデータを料金計算部126及び表

示制御部 1 2 7 に供給する。

[0119] 料金計算部 1 2 6 は、例えば、次式 (3) により、モデル利用者に課金する料金 C を計算する。

$$[0120] \quad C = C 1 \times \alpha_p \quad \dots (3)$$

[0121] C 1 は、0 より大きい所定の係数である。すなわち、料金 C は、予測精度 α_p に比例した金額となる。

[0122] 料金計算部 1 2 6 は、料金 C を示すデータを表示制御部 1 2 7 に供給する。

[0123] その後、処理はステップ S 1 2 に進む。

[0124] 一方、ステップ S 9 において、実績データに不正があると判定された場合、処理はステップ S 1 1 に進む。

[0125] ステップ S 1 1 において、サーバ 1 1 は、予測精度に基づく成果報酬と異なる体系で料金を計算する。

[0126] 具体的には、信頼度計算部 1 2 5 は、予測精度 α_p 、予測精度 α_q 、及び、実績データの不正の有無の判定結果を示すデータを料金計算部 1 2 6 及び表示制御部 1 2 7 に供給する。

[0127] 例えば、信頼度計算部 1 2 5 は、予測精度が 1 0 0 % の場合に上述した式 (3) により計算される金額より高い所定の金額に料金を設定する。

[0128] 或いは、例えば、信頼度計算部 1 2 5 は、モデル利用者の予測サービスの利用時間やデータ量に応じた金額を料金に設定する。

[0129] 或いは、例えば、信頼度計算部 1 2 5 は、次式 (4) により、料金 C を計算する。

$$[0130] \quad C = C 1 \times \alpha_q + C 2 (\alpha_q - \alpha_p) \quad \dots (4)$$

[0131] C 2 は、0 より大きい所定の係数である。式 (4) の第 1 項は、予測精度 α_f に近いと想定される予測精度 α_q に比例した金額となる。式 (4) の第 2 項は、予測精度 α_p が予測精度 α_q から乖離するほど高額になる懲罰的な金額となる。

[0132] 料金計算部 1 2 6 は、料金 C を示すデータを表示制御部 1 2 7 に供給する

- 。
- [0133] その後、処理はステップS 1 2に進む。
- [0134] ステップS 1 2において、サーバ1 1は、モデル提供者に対して料金と根拠を提示する。例えば、表示部1 0 3は、表示制御部1 2 7の制御の下に、図1 2の画面を表示する。
- [0135] 図1 2の画面には、ウィンドウ3 0 1、表示領域3 0 2、及び、請求ボタン3 0 3が上から順番に表示されている。
- [0136] ウィンドウ3 0 1内には、予測精度 α_p 及び予測精度 α_q を比較するための棒グラフが表示されている。また、その下に、実績データの信頼度R、及び、実績データの不正の有無の判定結果を示す数式が表示されている。さらに、その下に、不正検出パラメータ β の値、及び、不正検出パラメータ β を設定するためのスライダ3 1 1が表示されている。モデル提供者は、スライダ3 1 1を用いて不正検出パラメータ β を調整することができる。
- [0137] 表示領域3 0 2内には、モデル利用者に請求する料金、並びに、料金の根拠を説明するための文章及び数式が表示されている。この例では、予測対象ユーザに対する精度（予測精度 α_p ）と予測対象外ユーザに対する精度（予測精度 α_q ）との差が大きく、（予測対象ユーザの）実績データの信頼度が低い
ため、成果報酬と追加報酬を請求することが記載されている。
- [0138] 請求ボタン3 0 3は、モデル提供者がモデル利用者に請求する料金を確定する場合に押下される。
- [0139] ステップS 1 3において、信頼度計算部1 2 5は、不正検出パラメータ β が変更されたか否かを判定する。不正検出パラメータ β が変更されていないと判定された場合、処理はステップS 1 4に進む。
- [0140] ステップS 1 4において、料金計算部1 2 6は、料金を確定するか否かを判定する。料金を確定しないと判定された場合、処理はステップS 1 3に戻る。
- [0141] その後、ステップS 1 3において、不正検出パラメータ β が変更されたと判定されるか、ステップS 1 4において、料金を確定すると判定されるまで

、ステップS 1 3及びステップS 1 4の処理が繰り返し実行される。

[0142] 一方、ステップS 1 3において、例えば、モデル提供者が、サーバ1 1の入力部1 0 1を介して、図1 2の画面のスライダ3 1 1を操作した場合、信頼度計算部1 2 5は、不正検出パラメータ β が変更されたと判定し、処理はステップS 9に戻る。

[0143] その後、ステップS 1 4において、料金を確定すると判定されるまで、ステップS 9乃至ステップS 1 4の処理が繰り返し実行される。すなわち、スライダ3 1 1により変更された不正検出パラメータ β に基づいて、実績データの不正の有無を判定し、料金を計算し、提示する処理が繰り返し実行される。

[0144] 一方、ステップS 1 4において、例えば、モデル提供者が、サーバ1 1の入力部1 0 1を介して、図1 2の画面の請求ボタン3 0 3を押下した場合、料金計算部1 2 6は、料金を確定すると判定し、処理はステップS 1 5に進む。

[0145] ステップS 1 5において、サーバ1 1は、モデル利用者に対して料金と根拠を提示する。例えば、表示制御部1 2 7は、モデル利用者のクライアント1 2（実績データの送信元のクライアント1 2）において図1 3の画面を表示させるための表示制御データを生成する。表示制御部1 2 7は、通信部1 0 4を介して、表示制御データをクライアント1 2に送信する。

[0146] これに対して、クライアント1 2の表示制御部2 2 2は、ネットワーク1 3及び通信部2 0 4を介して、表示制御データを受信する。表示制御部2 2 2は、表示制御データに基づいて、表示部2 0 3を制御し、図1 3の画面を表示させる。

[0147] 図1 3の画面には、ウィンドウ3 5 1、表示領域3 5 2、支払いボタン3 5 3、及び、連絡ボタン3 5 4が表示されている。

[0148] ウィンドウ3 5 1内には、図1 2のウィンドウ3 0 1内からスライダ3 1 1を除いた情報が表示されている。

[0149] 表示領域3 5 2内には、図1 2の表示領域3 0 2内と同様の情報が表示さ

れている。

- [0150] 支払いボタン353は、モデル利用者が、提示された料金の支払いを行う場合に押下される。
- [0151] 連絡ボタン354は、モデル利用者が、提示された料金に対して問合せ等を行うためにモデル提供者に連絡する場合に押下される。
- [0152] その後、サービス提供処理は終了する。
- [0153] 以上のようにして、モデル利用者から提供される実績データ、特に、予測対象ユーザの実績データに対する信頼度を、容易かつ適正に計算することができる。また、信頼度に基づいて、モデル利用者から提供される実績データ、特に、予測対象ユーザの実績データの不正を確実に検出することができる。
- [0154] また、実績データの不正が確実に検出されることにより、モデル提供者及びモデル利用者に対して公平、かつ、適切な料金を請求することができる。例えば、実績データに不正がないと判定された場合、予測モデルの予測精度に応じた成果報酬により料金が請求される。一方、実績データに不正があると判定された場合、予測モデルの予測精度に応じた成果報酬とは異なる体系により計算された料金が計算される。これにより、モデル提供者の利益が確実に確保される。
- [0155] <<2. 変形例>>
- 以下、上述した本技術の実施の形態の変形例について説明する。
- [0156] 例えば、交差検証部123は、図10の交差検証用データに、図5の学習用データを追加して、交差検証を行うようにしてもよい。
- [0157] また、例えば、予測精度計算部124は、交差検証モデルの代わりに、予測モデルを用いて、予測対象ユーザに対する予測精度、及び、予測対象外ユーザに対する予測精度を計算するようにしてもよい。
- [0158] 例えば、予測精度計算部124は、予測対象外サンプルデータを予測モデルに入力することにより、予測対象外ユーザに対する予測確率を計算する。そして、予測精度計算部124は、予測部122により計算された予測対象

ユーザに対する予測確率、及び、予測対象ユーザの実績データに基づいて、予測対象ユーザに対する予測精度 α_p を計算する。また、予測精度計算部124は、予測対象外ユーザに対する予測確率、及び、予測対象外ユーザの実績データに基づいて、予測対象外ユーザに対する予測精度 α_q を計算する。

[0159] そして、信頼度計算部125は、例えば、上述した式(1)を用いて、予測精度 α_p 及び予測精度 α_q に基づいて、信頼度Rを計算する。

[0160] さらに、例えば、信頼度計算部125は、式(1)と異なる方法により、実績データの信頼度を計算するようにしてもよい。例えば、信頼度計算部125は、予測精度 α_p と予測精度 α_q との差に基づいて、信頼度を計算するようにしてもよい。

[0161] また、例えば、予測精度計算部124は、予測部122により予測された予測対象ユーザに対する予測確率と、予測対象ユーザの購入実績とに基づいて、実際の予測精度を計算し、料金計算部126が、式(3)において、予測精度 α_p の代わりに、実際の予測精度を用いて、料金Cを計算するようにしてもよい。

[0162] さらに、信頼度計算部125は、例えば、不正検出パラメータ β の初期値を、過去の各モデル利用者の実績に基づいて設定するようにしてもよい。例えば、信頼度計算部125は、あるモデル利用者に対して計算した予測精度のばらつきを考慮し、予測精度の標準偏差 σ に定数を乗じた値を不正検出パラメータ β に設定するようにしてもよい。この定数は、例えば、正規分布を仮定して、統計データの不正の発生確率が1%以下になるように設定される。

[0163] また、図12及び図13の画面は、その一例であり、任意に変更することが可能である。例えば、予測精度 α_p と予測精度 α_q 、及び、信頼度Rのうち一方のみ表示するようにしてもよい。また、例えば、モデル提供者用の画面と、モデル利用者用の画面とで、情報の内容や情報量を大きく変えるようにしてもよい。

[0164] さらに、以上では、予測対象となるサンプルが、サービスを利用するユー

ザ（人）である例を示したが、本技術は、人以外のサンプルに関する事象の予測を行う場合にも適用することができる。

[0165] また、例えば、サーバ 1 1 の機能の一部をクライアント 1 2 が実行するようにしてもよい。

[0166] <<3. その他>>

<コンピュータの構成例>

上述した一連の処理は、ハードウェアにより実行することもできるし、ソフトウェアにより実行することもできる。一連の処理をソフトウェアにより実行する場合には、そのソフトウェアを構成するプログラムが、コンピュータにインストールされる。ここで、コンピュータには、専用のハードウェアに組み込まれているコンピュータや、各種のプログラムをインストールすることで、各種の機能を実行することが可能な、例えば汎用のパーソナルコンピュータなどが含まれる。

[0167] 図 1 4 は、上述した一連の処理をプログラムにより実行するコンピュータのハードウェアの構成例を示すブロック図である。

[0168] コンピュータ 1 0 0 0 において、CPU (Central Processing Unit) 1 0 0 1, ROM (Read Only Memory) 1 0 0 2, RAM (Random Access Memory) 1 0 0 3 は、バス 1 0 0 4 により相互に接続されている。

[0169] バス 1 0 0 4 には、さらに、入出力インタフェース 1 0 0 5 が接続されている。入出力インタフェース 1 0 0 5 には、入力部 1 0 0 6、出力部 1 0 0 7、記録部 1 0 0 8、通信部 1 0 0 9、及びドライブ 1 0 1 0 が接続されている。

[0170] 入力部 1 0 0 6 は、入力スイッチ、ボタン、マイクロフォン、撮像素子などよりなる。出力部 1 0 0 7 は、ディスプレイ、スピーカなどよりなる。記録部 1 0 0 8 は、ハードディスクや不揮発性のメモリなどよりなる。通信部 1 0 0 9 は、ネットワークインタフェースなどよりなる。ドライブ 1 0 1 0 は、磁気ディスク、光ディスク、光磁気ディスク、又は半導体メモリなどのリムーバブルメディア 1 0 1 1 を駆動する。

- [0171] 以上のように構成されるコンピュータ1000では、CPU1001が、例えば、記録部1008に記録されているプログラムを、入出力インタフェース1005及びバス1004を介して、RAM1003にロードして実行することにより、上述した一連の処理が行われる。
- [0172] コンピュータ1000（CPU1001）が実行するプログラムは、例えば、パッケージメディア等としてのリムーバブルメディア1011に記録して提供することができる。また、プログラムは、ローカルエリアネットワーク、インターネット、デジタル衛星放送といった、有線または無線の伝送媒体を介して提供することができる。
- [0173] コンピュータ1000では、プログラムは、リムーバブルメディア1011をドライブ1010に装着することにより、入出力インタフェース1005を介して、記録部1008にインストールすることができる。また、プログラムは、有線または無線の伝送媒体を介して、通信部1009で受信し、記録部1008にインストールすることができる。その他、プログラムは、ROM1002や記録部1008に、あらかじめインストールしておくことができる。
- [0174] なお、コンピュータが実行するプログラムは、本明細書で説明する順序に沿って時系列に処理が行われるプログラムであっても良いし、並列に、あるいは呼び出しが行われたとき等の必要なタイミングで処理が行われるプログラムであっても良い。
- [0175] また、本明細書において、システムとは、複数の構成要素（装置、モジュール（部品）等）の集合を意味し、すべての構成要素が同一筐体中にあるか否かは問わない。したがって、別個の筐体に収納され、ネットワークを介して接続されている複数の装置、及び、1つの筐体の中に複数のモジュールが収納されている1つの装置は、いずれも、システムである。
- [0176] さらに、本技術の実施の形態は、上述した実施の形態に限定されるものではなく、本技術の要旨を逸脱しない範囲において種々の変更が可能である。
- [0177] 例えば、本技術は、1つの機能をネットワークを介して複数の装置で分担

、共同して処理するクラウドコンピューティングの構成をとることができる。
。

[0178] また、上述のフローチャートで説明した各ステップは、1つの装置で実行する他、複数の装置で分担して実行することができる。

[0179] さらに、1つのステップに複数の処理が含まれる場合には、その1つのステップに含まれる複数の処理は、1つの装置で実行する他、複数の装置で分担して実行することができる。

[0180] <構成の組み合わせ例>

本技術は、以下のような構成をとることもできる。

[0181] (1)

予測モデルの予測対象となる予測対象サンプルに対する実績値を示す第1の実績データに基づいて、前記予測対象サンプルに対する第1の予測精度を計算し、前記予測モデルの予測対象ではない予測対象外サンプルに対する実績値を示す第2の実績データに基づいて、前記予測対象外サンプルに対する第2の予測精度を計算する予測精度計算部と、

前記第1の予測精度及び前記第2の予測精度に基づいて、前記第1の実績データの信頼度を計算する信頼度計算部と

を備える情報処理装置。

(2)

前記予測精度計算部は、前記予測対象サンプルに関する予測対象サンプルデータ及び前記第1の実績データを含む第1のデータセットに基づいて、第1の予測精度を計算し、前記予測対象外サンプルに関する予測対象外サンプルデータ及び前記第2の実績データを含む第2のデータセットに基づいて、第2の予測精度を計算する

前記(1)に記載の情報処理装置。

(3)

前記第1のデータセット及び前記第2のデータセットを用いて交差検証を行う交差検証部を

さらに備え、

前記予測精度計算部は、前記交差検証の結果に基づいて、前記第1の予測精度及び前記第2の予測精度を計算する

前記(2)に記載の情報処理装置。

(4)

前記交差検証部は、交差検証に用いる訓練用データ及び検証用データに前記第1のデータセット内のデータ及び前記第2のデータセット内のデータが分散して含まれるように、前記第1のデータセット及び前記第2のデータセットを前記訓練用データ及び前記検証用データに分割する

前記(3)に記載の情報処理装置。

(5)

前記交差検証部は、前記予測モデルと同じ事象の予測を行う交差検証モデルを生成し、

前記予測精度計算部は、前記交差検証モデルを用いて、前記第1のデータセットに基づいて前記第1の予測精度を計算し、前記第2のデータセットに基づいて前記第2の予測精度を計算する

前記(3)又は(4)に記載の情報処理装置。

(6)

前記予測精度計算部は、前記予測モデルを用いて、前記第1のデータセットに基づいて前記第1の予測精度を計算し、前記第2のデータセットに基づいて前記第2の予測精度を計算する

前記(2)に記載の情報処理装置。

(7)

前記予測モデル及び前記予測対象サンプルデータに基づいて、所定の事象の予測を行う予測部を

さらに備える前記(2)乃至(6)のいずれかに記載の情報処理装置。

(8)

前記信頼度計算部は、前記第1の予測精度と前記第2の予測精度とを比較

することにより、前記信頼度を計算する

前記（１）乃至（７）のいずれかに記載の情報処理装置。

（９）

前記信頼度計算部は、前記信頼度に基づいて、前記第１の実績データの不正の有無を判定する

前記（８）に記載の情報処理装置。

（１０）

前記信頼度計算部は、前記信頼度と可変の閾値とを比較することにより、前記第１の実績データの不正の有無を判定する

前記（９）に記載の情報処理装置。

（１１）

前記第１の予測精度に基づいて、料金を計算する料金計算部を

さらに備える前記（１）乃至（１０）のいずれかに記載の情報処理装置。

（１２）

前記信頼度計算部は、前記信頼度に基づいて、前記第１の実績データの不正の有無を判定し、

前記料金計算部は、前記第１の実績データに不正がないと判定された場合、前記第１の予測精度に基づく成果報酬を前記料金として計算し、前記第１の実績データに不正があると判定された場合、前記第１の予測精度に基づく成果報酬とは異なる体系で前記料金を計算する

前記（１１）に記載の情報処理装置。

（１３）

前記料金及び前記料金の根拠の表示を制御する表示制御部を

さらに備える前記（１１）又は（１２）に記載の情報処理装置。

（１４）

前記表示制御部は、前記料金の根拠として、前記第１の予測精度と前記第２の予測精度、及び、前記信頼度のうち少なくとも１つを表示するように制御する

前記（１３）に記載の情報処理装置。

（１５）

前記表示制御部は、前記第１の実績データ及び前記第２の実績データの送信元である他の情報処理装置における前記料金及び前記料金の根拠の表示を制御する

前記（１３）又は（１４）に記載の情報処理装置。

（１６）

前記予測対象外サンプルは、前記予測モデルの学習対象となる学習対象サンプルを含む

前記（１）乃至（１５）のいずれかに記載の情報処理装置。

（１７）

前記学習対象サンプルに関する学習用データに基づいて、前記予測モデルの学習を行う学習部を

さらに備える前記（１６）に記載の情報処理装置。

（１８）

情報処理装置が、

予測モデルの予測対象となる予測対象サンプルに対する実績値を示す第１の実績データに基づいて、前記予測対象サンプルに対する第１の予測精度を計算し、

前記予測モデルの予測対象ではない予測対象外サンプルに対する実績値を示す第２の実績データに基づいて、前記予測対象外サンプルに対する第２の予測精度を計算し、

前記第１の予測精度及び前記第２の予測精度に基づいて、前記第１の実績データの信頼度を計算する

情報処理方法。

（１９）

予測モデルの予測対象となる予測対象サンプルに対する実績値を示す第１の実績データに基づいて、前記予測対象サンプルに対する第１の予測精度を

計算し、

前記予測モデルの予測対象ではない予測対象外サンプルに対する実績値を示す第2の実績データに基づいて、前記予測対象外サンプルに対する第2の予測精度を計算し、

前記第1の予測精度及び前記第2の予測精度に基づいて、前記第1の実績データの信頼度を計算する

処理をコンピュータに実行させるためのプログラム。

[0182] なお、本明細書に記載された効果はあくまで例示であって限定されるものではなく、他の効果があってもよい。

符号の説明

[0183] 1 情報処理システム, 11 サーバ, 12-1乃至12-n クライアント, 102 情報処理部, 103 表示部, 121 学習部, 122 予測部, 123 交差検証部, 124 予測精度計算部, 125 信頼度計算部, 126 料金計算部, 127 表示制御部, 202 情報処理部, 203 表示部, 221 演算処理部, 222 表示制御部

請求の範囲

- [請求項1] 予測モデルの予測対象となる予測対象サンプルに対する実績値を示す第1の実績データに基づいて、前記予測対象サンプルに対する第1の予測精度を計算し、前記予測モデルの予測対象ではない予測対象外サンプルに対する実績値を示す第2の実績データに基づいて、前記予測対象外サンプルに対する第2の予測精度を計算する予測精度計算部と、
- 前記第1の予測精度及び前記第2の予測精度に基づいて、前記第1の実績データの信頼度を計算する信頼度計算部と
- を備える情報処理装置。
- [請求項2] 前記予測精度計算部は、前記予測対象サンプルに関する予測対象サンプルデータ及び前記第1の実績データを含む第1のデータセットに基づいて、第1の予測精度を計算し、前記予測対象外サンプルに関する予測対象外サンプルデータ及び前記第2の実績データを含む第2のデータセットに基づいて、第2の予測精度を計算する
- 請求項1に記載の情報処理装置。
- [請求項3] 前記第1のデータセット及び前記第2のデータセットを用いて交差検証を行う交差検証部を
- さらに備え、
- 前記予測精度計算部は、前記交差検証の結果に基づいて、前記第1の予測精度及び前記第2の予測精度を計算する
- 請求項2に記載の情報処理装置。
- [請求項4] 前記交差検証部は、交差検証に用いる訓練用データ及び検証用データに前記第1のデータセット内のデータ及び前記第2のデータセット内のデータが分散して含まれるように、前記第1のデータセット及び前記第2のデータセットを前記訓練用データ及び前記検証用データに分割する
- 請求項3に記載の情報処理装置。

- [請求項5] 前記交差検証部は、前記予測モデルと同じ事象の予測を行う交差検証モデルを生成し、
- 前記予測精度計算部は、前記交差検証モデルを用いて、前記第1のデータセットに基づいて前記第1の予測精度を計算し、前記第2のデータセットに基づいて前記第2の予測精度を計算する
- 請求項3に記載の情報処理装置。
- [請求項6] 前記予測精度計算部は、前記予測モデルを用いて、前記第1のデータセットに基づいて前記第1の予測精度を計算し、前記第2のデータセットに基づいて前記第2の予測精度を計算する
- 請求項2に記載の情報処理装置。
- [請求項7] 前記予測モデル及び前記予測対象サンプルデータに基づいて、所定の事象の予測を行う予測部を
- さらに備える請求項2に記載の情報処理装置。
- [請求項8] 前記信頼度計算部は、前記第1の予測精度と前記第2の予測精度とを比較することにより、前記信頼度を計算する
- 請求項1に記載の情報処理装置。
- [請求項9] 前記信頼度計算部は、前記信頼度に基づいて、前記第1の実績データの不正の有無を判定する
- 請求項8に記載の情報処理装置。
- [請求項10] 前記信頼度計算部は、前記信頼度と可変の閾値とを比較することにより、前記第1の実績データの不正の有無を判定する
- 請求項9に記載の情報処理装置。
- [請求項11] 前記第1の予測精度に基づいて、料金を計算する料金計算部を
- さらに備える請求項1に記載の情報処理装置。
- [請求項12] 前記信頼度計算部は、前記信頼度に基づいて、前記第1の実績データの不正の有無を判定し、
- 前記料金計算部は、前記第1の実績データに不正がないと判定された場合、前記第1の予測精度に基づく成果報酬を前記料金として計算

し、前記第1の実績データに不正があると判定された場合、前記第1の予測精度に基づく成果報酬とは異なる体系で前記料金を計算する請求項11に記載の情報処理装置。

[請求項13] 前記料金及び前記料金の根拠の表示を制御する表示制御部をさらに備える請求項11に記載の情報処理装置。

[請求項14] 前記表示制御部は、前記料金の根拠として、前記第1の予測精度と前記第2の予測精度、及び、前記信頼度のうち少なくとも1つを表示するように制御する請求項13に記載の情報処理装置。

[請求項15] 前記表示制御部は、前記第1の実績データ及び前記第2の実績データの送信元である他の情報処理装置における前記料金及び前記料金の根拠の表示を制御する請求項13に記載の情報処理装置。

[請求項16] 前記予測対象外サンプルは、前記予測モデルの学習対象となる学習対象サンプルを含む請求項1に記載の情報処理装置。

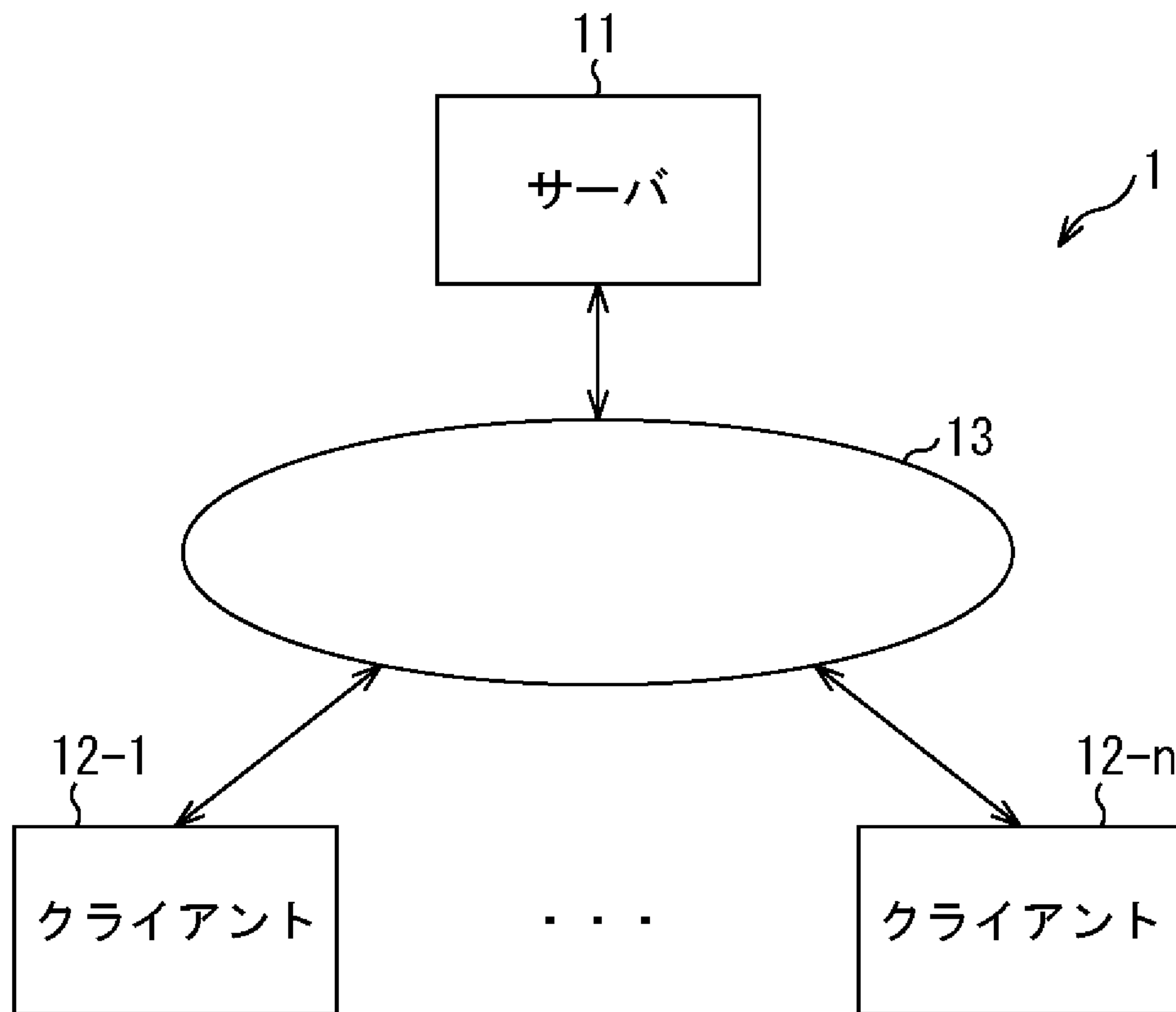
[請求項17] 前記学習対象サンプルに関する学習用データに基づいて、前記予測モデルの学習を行う学習部をさらに備える請求項16に記載の情報処理装置。

[請求項18] 情報処理装置が、
予測モデルの予測対象となる予測対象サンプルに対する実績値を示す第1の実績データに基づいて、前記予測対象サンプルに対する第1の予測精度を計算し、
前記予測モデルの予測対象ではない予測対象外サンプルに対する実績値を示す第2の実績データに基づいて、前記予測対象外サンプルに対する第2の予測精度を計算し、
前記第1の予測精度及び前記第2の予測精度に基づいて、前記第1の実績データの信頼度を計算する

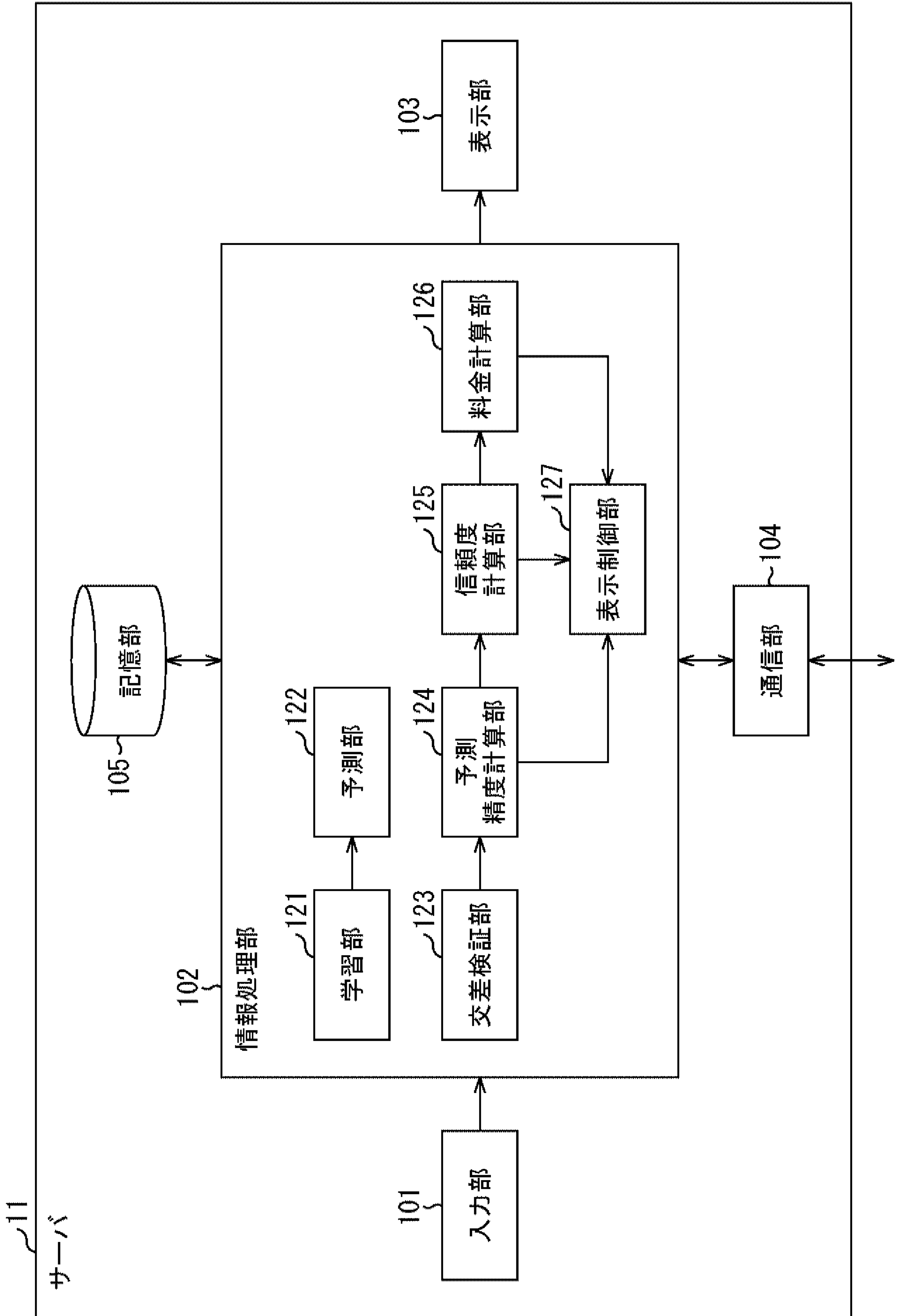
情報処理方法。

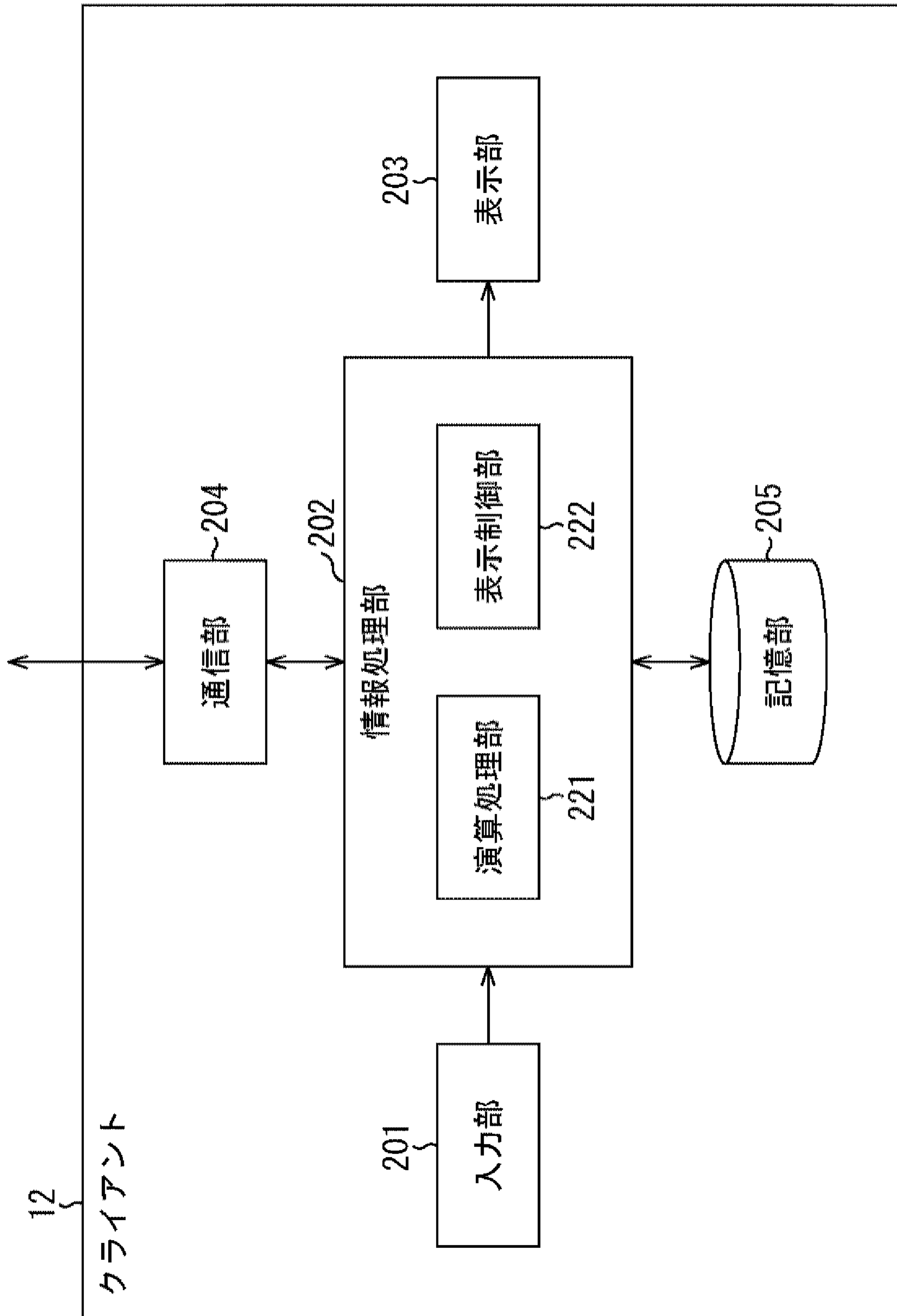
- [請求項19] 予測モデルの予測対象となる予測対象サンプルに対する実績値を示す第1の実績データに基づいて、前記予測対象サンプルに対する第1の予測精度を計算し、
- 前記予測モデルの予測対象ではない予測対象外サンプルに対する実績値を示す第2の実績データに基づいて、前記予測対象外サンプルに対する第2の予測精度を計算し、
- 前記第1の予測精度及び前記第2の予測精度に基づいて、前記第1の実績データの信頼度を計算する
- 処理をコンピュータに実行させるためのプログラム。

[図1]
FIG. 1

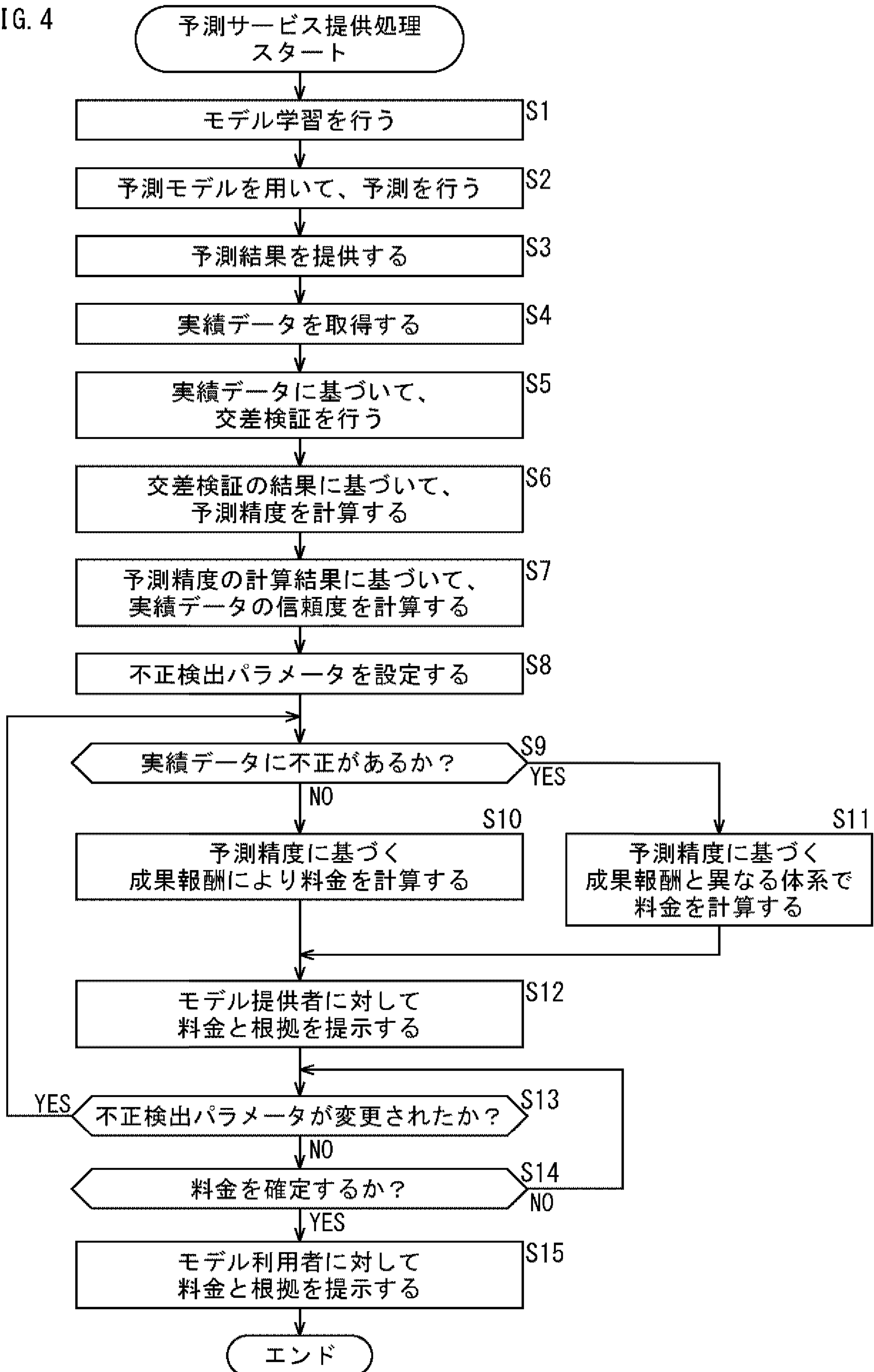


[図2]
FIG. 2



[図3]
FIG. 3

[図4]
FIG. 4



[図5]
FIG. 5

ID	性別	年齢	過去の 購入回数	購入実績
1001	M	24	3	0
1002	F	36	5	1
1003	F	53	14	1
1004	M	62	2	1
1005	M	33	1	0
.....				

[図6]
FIG. 6

ID	性別	年齢	過去の 購入回数
2001	M	51	2
2002	F	24	5
2003	M	34	4
2004	F	84	10
.....			
2234	M	71	1

[図7]
FIG. 7

ID	予測確率
2001	0.25
2002	0.87
2003	0.52
2004	0.16
.....	
2234	0.93

[図8]
FIG. 8

ID	購入実績
2001	0
2002	1
2003	1
2004	1
.....	
2234	0
3001	1
3002	0
3003	1
3004	1
.....	

} 予測対象

} 予測対象外

[図9]
FIG. 9

ID	性別	年齢	過去の 購入回数
3001	M	43	5
3002	F	18	3
3003	F	39	1
3004	M	28	7
.....			

[図10]
FIG. 10

ID	性別	年齢	過去の 購入回数	購入実績
2001	M	51	2	0
2002	F	24	5	1
2003	M	34	4	1
2004	F	84	10	1
.....				
2234	M	71	1	0
3001	M	43	5	1
3002	F	18	3	0
3003	F	39	1	1
3004	M	28	7	1
.....				

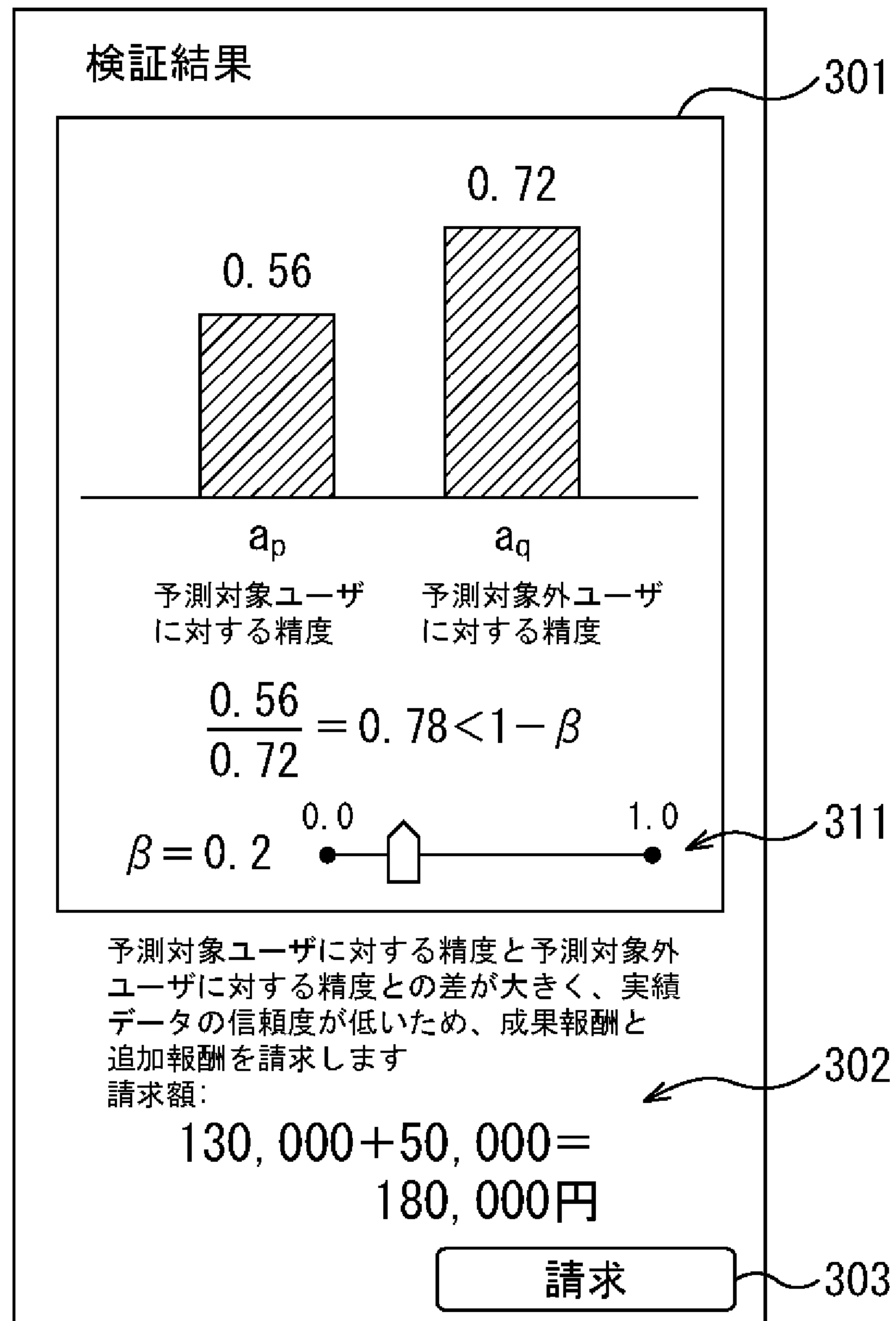
予測対象
予測対象外

[図11]
FIG. 11

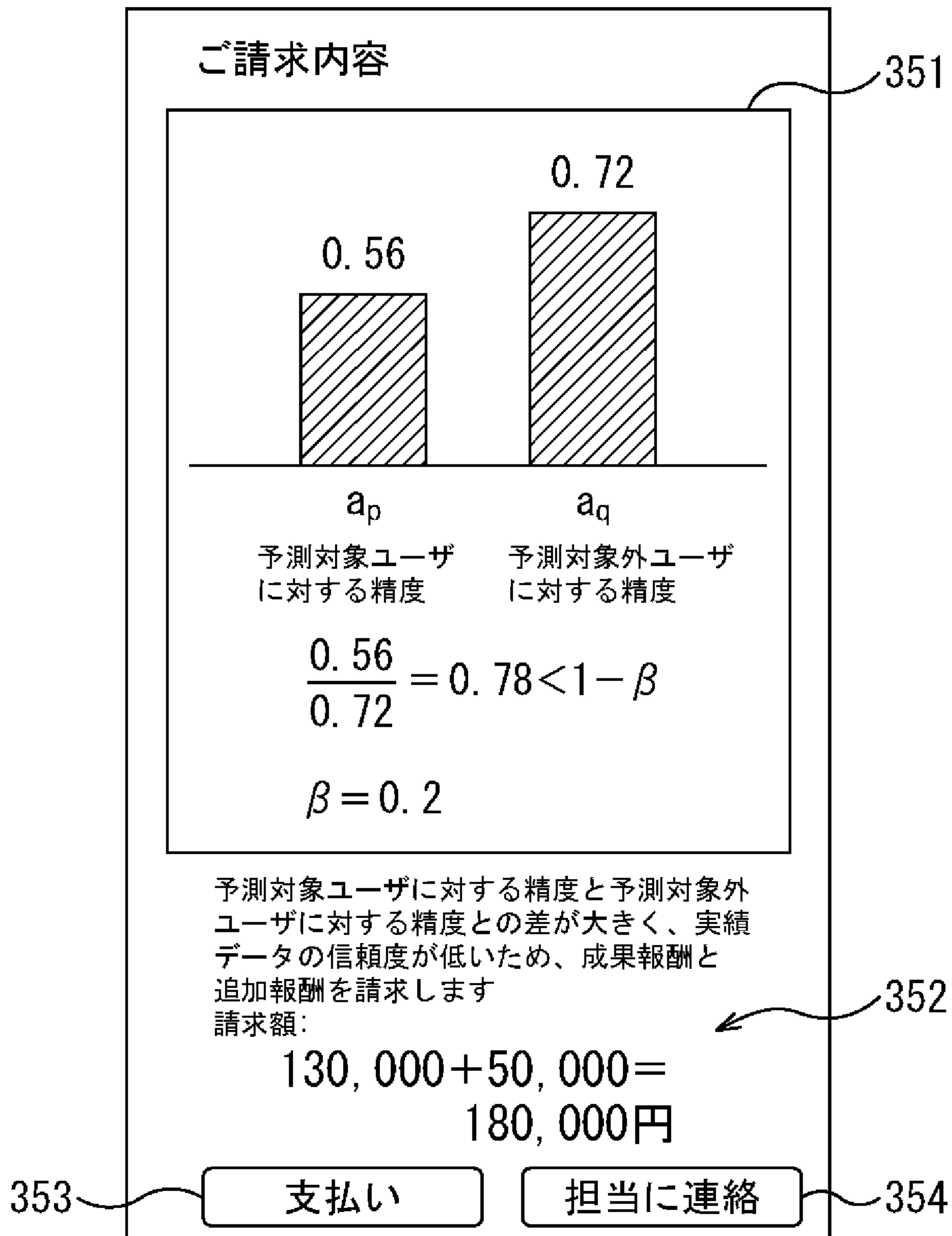
ID	性別	年齢	過去の 購入回数	購入実績	CV1	CV2	CV 予測確率
2001	M	51	2	0	Tr	Te	0.10
2002	F	24	5	1	Te	Tr	0.80
2003	M	34	4	1	Tr	Te	0.79
2004	F	84	10	1	Te	Tr	0.91
.....							
2234	M	71	1	0	Te	Tr	0.24
3001	M	43	5	1	Tr	Te	0.70
3002	F	18	3	0	Te	Tr	0.53
3003	F	39	1	1	Tr	Te	0.93
3004	M	28	7	1	Te	Tr	0.88
.....							

} 予測対象
} 予測対象外

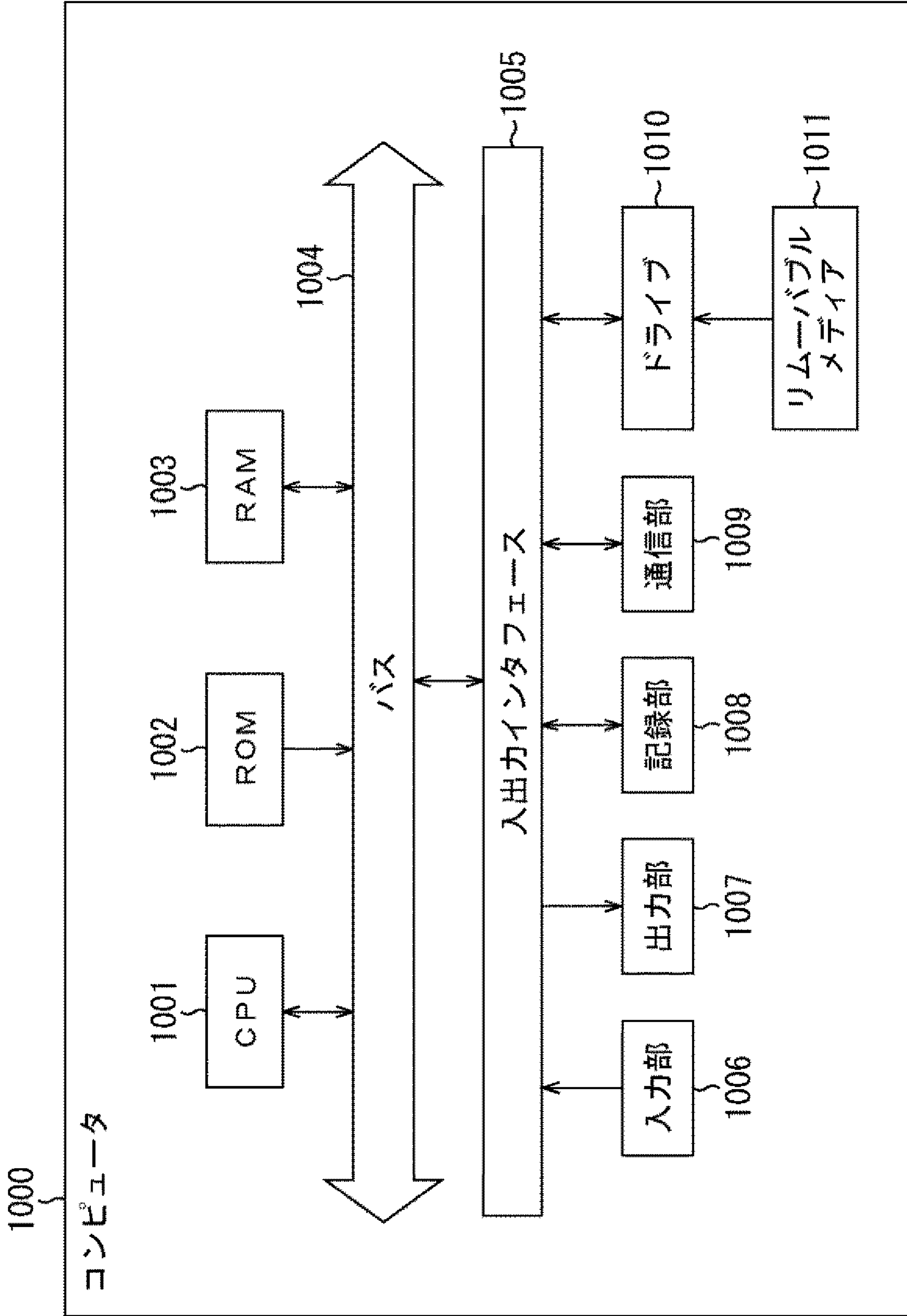
[図12]
FIG. 12



[図13]
FIG. 13



[図14]



INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2020/044987

<p>A. CLASSIFICATION OF SUBJECT MATTER G06Q 10/04 (2012.01) i FI: G06Q10/04 According to International Patent Classification (IPC) or to both national classification and IPC</p>										
<p>B. FIELDS SEARCHED</p>										
<p>Minimum documentation searched (classification system followed by classification symbols) G06Q10/04</p>										
<p>Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched</p> <table style="width:100%; border:none;"> <tr> <td style="width:80%;">Published examined utility model applications of Japan</td> <td style="width:20%; text-align:right;">1922-1996</td> </tr> <tr> <td>Published unexamined utility model applications of Japan</td> <td style="text-align:right;">1971-2021</td> </tr> <tr> <td>Registered utility model specifications of Japan</td> <td style="text-align:right;">1996-2021</td> </tr> <tr> <td>Published registered utility model applications of Japan</td> <td style="text-align:right;">1994-2021</td> </tr> </table>			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
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									
<p>Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)</p>										
<p>C. DOCUMENTS CONSIDERED TO BE RELEVANT</p>										
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.								
A	JP 2019-087027 A (TOSHIBA CORP.) 06 June 2019 (2019-06-06) entire text, all drawings	1-19								
A	JP 2008-123165 A (MIZEN KK) 29 May 2008 (2008-05-29) entire text, all drawings	1-19								
A	JP 2004-139196 A (HITACHI, LTD.) 13 May 2004 (2004-05-13) entire text, all drawings	1-19								
<p><input type="checkbox"/> Further documents are listed in the continuation of Box C. <input checked="" type="checkbox"/> See patent family annex.</p>										
<p>* Special categories of cited documents:</p> <table style="width:100%; border:none;"> <tr> <td style="width:50%;"> <p>“A” document defining the general state of the art which is not considered to be of particular relevance</p> <p>“E” earlier application or patent but published on or after the international filing date</p> <p>“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)</p> <p>“O” document referring to an oral disclosure, use, exhibition or other means</p> <p>“P” document published prior to the international filing date but later than the priority date claimed</p> </td> <td style="width:50%;"> <p>“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</p> <p>“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</p> <p>“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</p> <p>“&” document member of the same patent family</p> </td> </tr> </table>			<p>“A” document defining the general state of the art which is not considered to be of particular relevance</p> <p>“E” earlier application or patent but published on or after the international filing date</p> <p>“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)</p> <p>“O” document referring to an oral disclosure, use, exhibition or other means</p> <p>“P” document published prior to the international filing date but later than the priority date claimed</p>	<p>“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</p> <p>“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</p> <p>“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</p> <p>“&” document member of the same patent family</p>						
<p>“A” document defining the general state of the art which is not considered to be of particular relevance</p> <p>“E” earlier application or patent but published on or after the international filing date</p> <p>“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)</p> <p>“O” document referring to an oral disclosure, use, exhibition or other means</p> <p>“P” document published prior to the international filing date but later than the priority date claimed</p>	<p>“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</p> <p>“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</p> <p>“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</p> <p>“&” document member of the same patent family</p>									
<p>Date of the actual completion of the international search 28 January 2021 (28.01.2021)</p>		<p>Date of mailing of the international search report 09 February 2021 (09.02.2021)</p>								
<p>Name and mailing address of the ISA/ Japan Patent Office 3-4-3, Kasumigaseki, Chiyoda-ku, Tokyo 100-8915, Japan</p>		<p>Authorized officer</p> <p>Telephone No.</p>								

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application no.
PCT/JP2020/044987

Patent Documents referred in the Report	Publication Date	Patent Family	Publication Date
JP 2019-087027 A	06 Jun. 2019	US 2019/0139059 A1	
JP 2008-123165 A	29 May 2008	(Family: none)	
JP 2004-139196 A	13 May 2004	(Family: none)	

A. 発明の属する分野の分類 (国際特許分類 (IPC)) G06Q 10/04(2012.01)i FI: G06Q10/04		
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)) G06Q10/04		
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)		
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
A	JP 2019-087027 A (株式会社東芝) 06.06.2019 (2019 - 06 - 06) 全文全図	1-19
A	JP 2008-123165 A (有限会社 みぜん) 29.05.2008 (2008 - 05 - 29) 全文全図	1-19
A	JP 2004-139196 A (株式会社日立製作所) 13.05.2004 (2004 - 05 - 13) 全文全図	1-19
<input type="checkbox"/> C欄の続きにも文献が列挙されている。 <input checked="" type="checkbox"/> 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	
28.01.2021	09.02.2021	
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) ▲高▼瀬 健太郎 5E 3865 電話番号 03-3581-1101 内線 3521	

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/JP2020/044987

引用文献			公表日	パテントファミリー文献			公表日
JP	2019-087027	A	06.06.2019	US	2019/0139059	A1	
JP	2008-123165	A	29.05.2008	(ファミリーなし)			
JP	2004-139196	A	13.05.2004	(ファミリーなし)			