

(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2月27日(27.02.2020)



(10) 国際公開番号

WO 2020/040253 A1

- (51) 国際特許分類:
G06Q 10/04 (2012.01) *G16H 50/30* (2018.01)
G06N 20/00 (2019.01)
- (21) 国際出願番号: PCT/JP2019/032843
- (22) 国際出願日: 2019年8月22日(22.08.2019)
- (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 2018-157718 2018年8月24日(24.08.2018) JP
- (71) 出願人:株式会社 N T T ドコモ(NTT DOCOMO, INC.) [JP/JP]; 〒1006150 東京都千代田区永田町2丁目11番1号 Tokyo (JP).
- (72) 発明者:伊藤 拓(ITO, Taku); 〒1006150 東京都千代田区永田町2丁目11番1号 山王パークタワー 株式会社 N T T ドコモ 知的財産部内 Tokyo (JP). 落合 桂一(OCHIAI, Keiichi); 〒1006150 東京都千代田区永田町2丁目11番1号 山王パークタワー 株式会社 N T T ドコモ 知的財産部内 Tokyo (JP). 深澤 佑介(FUKAZAWA, Yusuke); 〒1006150 東京都千代田区永田町2丁目11番1号 山王パークタワー 株式会社 N T T ドコモ 知的財産部内 Tokyo (JP).
- (74) 代理人:伊東 忠重, 外(ITO, Tadashige et al.); 〒1000005 東京都千代田区丸の内二丁目1番1号 丸の内 M Y P L A Z A (明治安田生命ビル) 16階 Tokyo (JP).

(54) Title: PREDICTION INTERPRETATION DEVICE AND PREDICTION INTERPRETATION METHOD

(54) 発明の名称: 予測解釈装置、及び予測解釈方法

〔図1〕

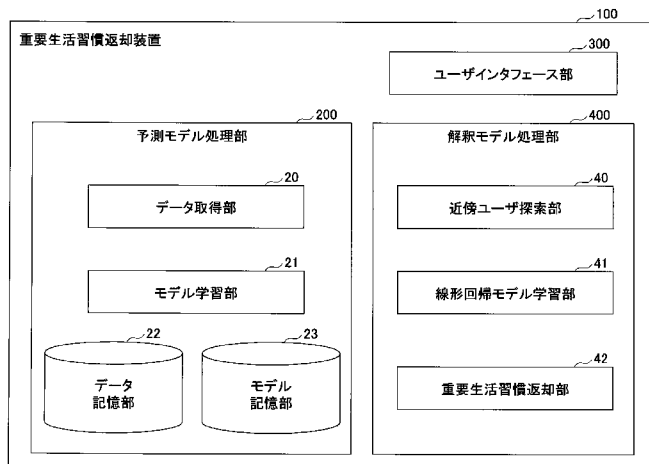


FIG. 1:

- 20 Data acquisition unit
- 21 Model learning unit
- 22 Data storage unit
- 23 Model storage unit
- 40 Neighborhood user search unit
- 41 Linear regression model learning unit
- 42 Important lifestyle return unit
- 100 Important lifestyle return device
- 200 Prediction model processing unit
- 300 User interface section
- 400 Interpretation model processing unit

(57) Abstract: This prediction interpretation device comprises: a data storage unit which stores data about a plurality of users; a model storage unit which stores a prediction model learned from the data about all of the plurality of users; a neighborhood user search unit which extracts, from the data storage unit, neighborhood users around a target user; a linear regression model learning unit which learns, for the neighborhood users, a linear regression model approximated to the prediction model; and an interpretation result output unit which outputs an interpretation result of prediction for the target user on the basis of the partial regression coefficient of the linear regression model. The neighborhood user search unit extracts the neighborhood users by narrowing down, on the basis of a prediction direction of the target user by the prediction model, neighborhood user candidates extracted on the basis of a distance between the users.

(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, 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, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, 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: 予測解釈装置において、複数ユーザのデータを格納するデータ記憶部と、前記複数ユーザ全体のデータから学習された予測モデルを格納するモデル記憶部と、対象ユーザに対する近傍ユーザを前記データ記憶部から抽出する近傍ユーザ探索部と、前記近傍ユーザを対象として、前記予測モデルに近似させた線形回帰モデルを学習する線形回帰モデル学習部と、前記線形回帰モデルの偏回帰係数に基づいて前記対象ユーザに対する予測の解釈結果を出力する解釈結果出力部と、を備え、前記近傍ユーザ探索部は、ユーザ間の距離に基づき抽出された近傍ユーザ候補を、前記予測モデルによる前記対象ユーザの予測方向に基づいて絞り込むことにより、前記近傍ユーザを抽出する。

明 細 書

発明の名称： 予測解釈装置、及び予測解釈方法

技術分野

[0001] 本発明は、機械学習に基づく予測結果の解釈を行う技術に関連するものである。本技術の適用分野の一例として、健康予測を高い精度で行いつつ、健康に影響を与える生活習慣を独立して算出する技術がある。

背景技術

[0002] 近年、機械学習によって構築した予測モデルを用いて、入力データから結果を予測する技術が用いられている。

[0003] また、得られた予測結果に関して、入力データにおけるどの要因が予測結果に大きく影響しているか等、予測結果の解釈を行うための技術も用いられている。そのような予測結果の解釈を可能とする技術として、例えばLIMEと呼ばれる技術がある。LIMEでは、機械学習における予測モデルと、予測の解釈用モデルとを独立して算出することで、機械学習における予測結果の解釈を可能としている。

[0004] また、予測結果の解釈を可能とする技術として、ランダム森の重要度を利用する技術（例えば特許文献1）、決定木と回帰を組み合わせることでモデルを学習する異種混合学習（例えば非特許文献1）等もある。

先行技術文献

特許文献

[0005] 特許文献1：特開2009-057337号公報

非特許文献

[0006] 非特許文献1：Riki Eto, Ryohei Fujimaki, Satoshi Morinaga, Hiroshi Tamano; Proceedings of the Seventeenth International Conference on Artificial Intelligence and Statistics, PMLR 33: 238-246, 2014.

発明の概要

発明が解決しようとする課題

[0007] しかし、上述したLIMEでは、予測を高い精度で行うことができて、直観に反する解釈を与える場合があり、解釈精度が低いという問題があった。また、特許文献1及び非特許文献2に開示された技術では、予測精度を上げることと解釈精度を上げることがトレードオフの関係にあり、両方を高い精度で実施することが難しいという問題があった。

[0008] 本発明は上記の点に鑑みてなされたものであり、機械学習による予測と、予測の解釈とを高い精度で行うことを可能とする技術を提供することを目的とする。

課題を解決するための手段

[0009] 開示の技術によれば、複数ユーザのデータを格納するデータ記憶部と、前記複数ユーザ全体のデータから学習された予測モデルを格納するモデル記憶部と、

対象ユーザに対する近傍ユーザを前記データ記憶部から抽出する近傍ユーザ探索部と、

前記近傍ユーザを対象として、前記予測モデルに近似させた線形回帰モデルを学習する線形回帰モデル学習部と、

前記線形回帰モデルの偏回帰係数に基づいて前記対象ユーザに対する予測の解釈結果を出力する解釈結果出力部と、を備え、

前記近傍ユーザ探索部は、ユーザ間の距離に基づき抽出された近傍ユーザ候補を、前記予測モデルによる前記対象ユーザの予測方向に基づいて絞り込むことにより、前記近傍ユーザを抽出する

予測解釈装置が提供される。

発明の効果

[0010] 開示の技術によれば、機械学習による予測と、予測の解釈とを高い精度で行うことを可能とする技術が提供される。

図面の簡単な説明

[0011] [図1]本発明の実施の形態にかかる重要生活習慣返却装置100の機能構成図である。

[図2]重要生活習慣返却装置 100 のハードウェア構成の一例を示す図である。

[図3]重要生活習慣返却装置 100 の動作を説明するためのフローチャートである。

[図4]重要生活習慣返却装置 100 の動作を説明するためのフローチャートである。

[図5]近傍ユーザの抽出を説明するための図である。

[図6]重要生活習慣返却装置 100 により出力される画面例を示す図である。

[図7]重要生活習慣返却装置 100 により出力される画面例を示す図である。

発明を実施するための形態

[0012] 以下、図面を参照して本発明の実施の形態（本実施の形態）を説明する。
なお、以下で説明する実施の形態は一例に過ぎず、本発明が適用される実施の形態は、以下の実施の形態に限られるわけではない。

[0013] 以下では本発明の一実施形態として、健康リスクに関わる重要生活習慣を提示（返却）する重要生活習慣返却装置 100 について説明するが、本発明の適用分野はこれに限られるわけではなく、本発明は様々な分野において適用可能である。

[0014] （重要生活習慣返却装置 100 の機能構成例）

図 1 は、本実施の形態における重要生活習慣返却装置 100 の機能構成図である。図 1 に示すように、本実施の形態における重要生活習慣返却装置 100 は、予測モデル処理部 200 と、ユーザインタフェース部 300 と、解釈モデル処理部 400 とを有する。

[0015] 予測モデル処理部 200 は、実際にユーザがユーザ端末からアクセスする前に学習を行う処理部であり、複数ユーザの全体のデータから予測モデルを構築する。解釈モデル処理部 400 は、生活習慣の改善案の提示を求めるユーザ（対象ユーザ）の近傍ユーザのデータを用いた解釈用モデルの作成等を行う。

[0016] 本実施の形態では、重要生活習慣返却装置 100 とユーザ端末とがネット

ワークを介して接続されることを想定しており、ユーザインタフェース部 300 は、ネットワークを介してユーザ端末に画面（Web 画面等）を表示したり、ユーザ端末からの入力情報を受け付ける等の処理を実施する。例えば、ユーザインタフェース部 300 は、生活習慣の改善案の提示を求めるユーザのユーザ端末からのアクセスを受け付ける。なお、ユーザが直接に重要生活習慣返却装置 100 を操作する構成であってもよい。その場合、ユーザインタフェース部 300 は、例えば、ディスプレイ及びキーボード、入出力操作可能なタッチパネル等である。以下、予測モデル処理部 200 と解釈モデル処理部 300 をより詳細に説明する。

[0017] 図 1 に示すように、予測モデル処理部 200 は、データ取得部 20 と、モデル学習部 21 と、データ記憶部 22 と、モデル記憶部 23 とを有する。

[0018] データ取得部 20 は、全ユーザの健康診断データ及び生活習慣の問診結果データを取得し、取得したデータをデータ記憶部 22 に格納する。健康診断データ及び生活習慣の問診結果データは、例えば、これらを蓄積した所定のサーバからネットワークを介して取得される。また、取得されるデータには、健康診断データ及び生活習慣の問診結果データと紐付く形で各ユーザの属性等も含まれる。なお、健康診断データ及び生活習慣の問診結果データを、「属性」と見なしてもよい。

[0019] モデル学習部 21 は、データ取得部 20 により取得されたデータをデータ記憶部 22 から読み出し、当該データを用いてユーザ全体における健康リスクの予測モデルを学習し、学習された予測モデルをモデル記憶部 23 に格納する。

[0020] データ記憶部 22 及びモデル記憶部 23 はそれぞれ上述したデータを格納する。なお、データ記憶部 22 及びモデル記憶部 23 は、予測モデル処理部 200 の外部にあってもよい。また、データ記憶部 22 及びモデル記憶部 23 が、重要生活習慣返却装置 100 の外部のネットワーク上にあってもよい。

[0021] 解釈モデル処理部 400 は、近傍ユーザ探索部 40 と、線形回帰モデル学

習部 4 1 と、重要生活習慣返却部 4 2 とを有する。

[0022] 近傍ユーザ探索部 4 0 は、データ記憶部 2 2 を参照することにより、対象ユーザに近い属性を持ち、なおかつ、対象ユーザの予測モデルにおける予測結果と大きく外れていないユーザを近傍ユーザとして抽出する。

[0023] 線形回帰モデル学習部 4 1 は、近傍ユーザ探索部 4 0 により抽出された近傍ユーザのみについて、予測モデル処理部 2 0 0 で学習した予測モデルに近似する線形回帰モデルを学習し、学習した線形回帰モデルをモデル記憶部 2 3 に格納する。この線形回帰モデルは解釈用の予測モデルであり、これを解釈用モデルと呼ぶことができる。

[0024] 重要生活習慣返却部 4 2 は、線形回帰モデル学習部 4 1 で学習された線形回帰モデルの偏回帰係数をもとに、対象ユーザにとって改善すべき重要な生活習慣をユーザインタフェース部 3 0 0 を介してユーザに返却する。

[0025] (ハードウェア構成例)

図 1 に示した重要生活習慣返却装置 1 0 0 のブロック図は、機能単位のブロックを示している。これらの機能ブロック（構成部）は、例えば、ハードウェア及びソフトウェアの組み合わせによって実現される。

[0026] また、重要生活習慣返却装置 1 0 0 は、本実施の形態に係る処理を行うコンピュータで実現してもよい。図 2 は、重要生活習慣返却装置 1 0 0 として使用されるコンピュータのハードウェア構成の一例を示す図である。図 2 に示すように、当該コンピュータは、プロセッサ 1 0 0 1、メモリ 1 0 0 2、ストレージ 1 0 0 3、通信装置 1 0 0 4、入力装置 1 0 0 5、出力装置 1 0 0 6、バス 1 0 0 7 を含む。

[0027] 重要生活習慣返却装置 1 0 0 における各機能は、プロセッサ 1 0 0 1、メモリ 1 0 0 2 等のハードウェア上に所定のソフトウェア（プログラム）を読み込ませることで、プロセッサ 1 0 0 1 が演算を行い、通信装置 1 0 0 4 による通信、メモリ 1 0 0 2 及びストレージ 1 0 0 3 におけるデータの読み出し及び書き込みを制御することで実現される。

[0028] プロセッサ 1 0 0 1 は、例えば、オペレーティングシステムを動作させて

コンピュータ全体を制御する。プロセッサ１００１は、周辺装置とのインターフェース、制御装置、演算装置、レジスタ等を含む中央処理装置（CPU：Central Processing Unit）で構成されてもよい。

[0029] また、プロセッサ１００１は、プログラム（プログラムコード）、ソフトウェアモジュール又はデータを、ストレージ１００３又は通信装置１００４からメモリ１００２に読み出し、これらに従って各種の処理を実行する。プログラムとしては、本実施の形態で説明する動作の少なくとも一部をコンピュータに実行させるプログラムが用いられる。なお、プログラムは、電気通信回線を介してネットワークから送信されても良い。

[0030] メモリ１００２は、コンピュータ読み取り可能な記録媒体であり、例えば、ROM（Read Only Memory）、EPROM（Erasable Programmable ROM）、EEPROM（Electrically Erasable Programmable ROM）、RAM（Random Access Memory）等の少なくとも１つで構成されてもよい。メモリ１００２は、レジスタ、キャッシュ、メインメモリ（主記憶装置）等と呼ばれてもよい。メモリ１００２は、本発明の実施の形態に係る処理を実施するために実行可能なプログラム（プログラムコード）、ソフトウェアモジュール等を保存することができる。

[0031] ストレージ１００３は、コンピュータ読み取り可能な記録媒体であり、例えば、CD-ROM（Compact Disc ROM）等の光ディスク、ハードディスクドライブ、フレキシブルディスク、光磁気ディスク（例えば、コンパクトディスク、デジタル多用途ディスク、Blu-ray（登録商標）ディスク）、スマートカード、フラッシュメモリ（例えば、カード、スティック、キードライブ）、フロッピー（登録商標）ディスク、磁気ストリップ等の少なくとも１つで構成されてもよい。ストレージ１００３は、補助記憶装置と呼ばれてもよい。上述の記録媒体は、例えば、メモリ１００２又はストレージ１００３を含むデータベース、サーバその他の適切な媒体で

あってもよい。

[0032] 通信装置1004は、有線又は無線ネットワークを介してコンピュータ間の通信を行うためのハードウェア（送受信デバイス）であり、例えばネットワークデバイス、ネットワークコントローラ、ネットワークカード、通信モジュール等ともいう。

[0033] 入力装置1005は、外部からの入力を受け付ける入力デバイス（例えば、キーボード、マウス、マイクロフォン、スイッチ、ボタン、センサ等）である。出力装置1006は、外部への出力を実施する出力デバイス（例えば、ディスプレイ、スピーカー、LEDランプ等）である。なお、入力装置1005及び出力装置1006は、一体となった構成（例えば、タッチパネル）であってもよい。

[0034] また、プロセッサ1001及びメモリ1002等の各装置は、情報を通信するためのバス1007で接続される。バス1007は、単一のバスで構成されてもよいし、装置間で異なるバスで構成されてもよい。

[0035] また、重要生活習慣返却装置100は、マイクロプロセッサ、デジタル信号プロセッサ（DSP: Digital Signal Processor）、ASIC（Application Specific Integrated Circuit）、PLD（Programmable Logic Device）、FPGA（Field Programmable Gate Array）等のハードウェアを含んで構成されてもよく、当該ハードウェアにより、各機能ブロックの一部又は全てが実現されてもよい。例えば、プロセッサ1001は、これらのハードウェアの少なくとも1つで実装されてもよい。

[0036] （重要生活習慣返却装置100の動作）

次に、図1に示した機能構成を備える重要生活習慣返却装置100の動作を図3及び図4のフローチャートを参照して詳細に説明する。以下の動作において、データ取得部20により取得したユーザ全体の健康診断データ及び生活習慣の問診データは既にデータ記憶部22に格納されているものとする

。

[0037] まず、図3を参照して説明する。図3におけるS101の処理は、対象ユーザからのアクセスがある前に予め行う処理であり、S102は実際に重要生活習慣返却装置100へ対象ユーザからアクセスがあった場合に行う処理である。図4は、S102の詳細内容を示す。

[0038] まず、S101において、モデル学習部21が、データ記憶部22からユーザ全体の健康診断データ及び生活習慣の問診データを読み出し、これらのデータを使用して、XGBoostを一例とする機械学習ソフトウェアにより予測モデルを学習する。この予測モデルにより、例えば、生活習慣のデータに基づいて、収縮時血圧を一例とする健康リスク値について、1年後、2年後、3年後の結果値を予測することができる。

[0039] 続いて、ユーザインタフェース部300を介してあるユーザ（対象ユーザ）のユーザ端末からアクセスを受けると、S102において、解釈モデル処理部400が、対象ユーザに対する近傍のユーザのみを学習対象とする解釈用モデルを学習し、重要生活習慣返却部42が、解釈用モデルに基づいて対象ユーザに対して重要な生活習慣を提示する。なお、「重要な生活習慣」は、予測に対する解釈結果の例である。

[0040] （S102の詳細）

以下、S102における解釈用モデルの学習処理、及び重要な生活習慣の提示処理の詳細を、図4のフローチャート、及び図5～7を参照して説明する。

[0041] ＜S201：近傍ユーザの抽出＞

まず、図4のS201において、ユーザインタフェース部300が対象ユーザからのリクエストを受け付けた後、近傍ユーザ探索部40が、対象ユーザと健康状態及び生活習慣が近いユーザを、データ記憶部22に記憶された全ユーザから探索し、回帰モデルを構築する際に用いる近傍ユーザの情報を線形回帰モデル学習部41へ引き渡す。

[0042] なお、対象ユーザは、S101において学習に用いたユーザに含まれてい

なくてもよい。例えば、対象ユーザがユーザインタフェース部300を介して自身の健康状態及び生活習慣を入力し、これら健康状態及び生活習慣を用いて近傍ユーザを探索することができる。

[0043] 近傍ユーザ探索部40が行う探索法についてより具体的に説明する。

[0044] 対象ユーザの状態ベクトル（例えば、BMI、収縮時血圧等の健康リスク値及び運動時間や朝食の有無等の生活習慣値を表したベクトル）を x とし、近傍ユーザの状態ベクトルを x' としたとき、近傍ユーザとの距離はユークリッド距離 $D(x, x')$ で表される。 $D(x, x')$ が他と比べて小さい所定数のユーザを、対象ユーザの近傍ユーザとする。なお、ここでの近傍ユーザは最終的な近傍ユーザではないので、近傍ユーザ候補と称してもよい。

[0045] 続いて、近傍ユーザ探索部40は近似結果の妥当性に基づく近傍ユーザの絞り込みを行う。これにより、生活習慣を改善したにも関わらず健康リスク値が上がってしまうといった誤予測を防ぐ効果がある。

[0046] 具体的には、モデル学習部21で学習された予測モデルを f とおいたとき、予測モデルの微分係数

[0047] [数1]

$$\frac{\partial f}{\partial x}$$

と、対象ユーザと近傍ユーザ間の健康リスク値の傾き

[0048] [数2]

$$\frac{f(x') - f(x)}{|x' - x|}$$

を計算する。

[0049] 「数1」で示される微分係数は、対象ユーザの状態を微小変化させたときの f の変化量を表す。簡単のために、対象ユーザの状態 x が1次元であるとする、 x と $f(x)$ とを2次元グラフで表したときに、「数1」で示される微分係数は、 f のグラフの点 $(x, f(x))$ における接線の傾きを表す。なお、当該微分係数を予測方向と称してもよい。

[0050] また、「数2」は、近傍ユーザ候補についての状態 x' と予測モデル f による予測結果 $f(x')$ とに対応する点と、対象ユーザについての状態 x と予測モデル f による予測結果 $f(x)$ とに対応する点とを結んだ直線の傾きを表す。

[0051] 近傍ユーザ探索部40は、対象ユーザと近傍ユーザ間の健康リスク値の傾きと、微分係数との積を計算し、それが負である場合にはそのユーザを近傍ユーザから除外する。すなわち、

[0052] [数3]

$$\frac{\partial f}{\partial x} \cdot \frac{f(x') - f(x)}{|x' - x|}$$

が負となるユーザ (x') を近傍ユーザから除外する。このことは、対象ユーザの状態が変化した場合の予測値の変化傾向（予測方向）から大きくはずれた予測値を持つ近傍ユーザを、最終的な近傍ユーザから除外することを意味する。この意味を図5を参照してより具体的に説明する。

[0053] 図5において、「A」で示される曲線が予測モデル f を表し、Bで示される点線が対象ユーザと近傍ユーザとの間での健康リスク値の傾きを表している。対象ユーザについて、現状の状態の点における予測モデル f の接線の傾きが右上を向くので、予測モデル f では飲酒量が増えるほどBMIが上ると予測されている。

[0054] 一方、飲酒量が少ないが極端にBMIの高い予測結果（ $f(x')$ ）が得られた近傍ユーザがいた場合、近似の回帰モデルを生成した際に誤予測となる可能性がある。そのため、Bの点線で示すように対象ユーザと近傍ユーザとの間の傾きを求め、予測モデル f における微分係数と傾きとの積が負である場合には、外れ値であるとして、当該近傍ユーザを最終的な近傍ユーザから除外することとしている。

[0055] 図5に示す g （Cで示す直線）は、最終的な近傍ユーザのデータに基づく計算により得られた解釈用モデルである。これについては、下記のS202で説明する。

[0056] <S202：線形回帰モデル g の学習>

図4のS202では、S201で抽出された近傍ユーザ群 $X' = \{x' \in X \mid \dots\}$ において、線形回帰モデル学習部41が、モデル記憶部23に記憶されている学習済み予測モデル f とフィッティングするような線形回帰モデル g を解釈用モデルとして学習する。具体的には、

[0057] [数4]

$$L(f, g, \pi_x) = \sum_{x' \in X'} \pi_x(x') (f(x') - g(x'))^2$$

で定義される損失関数 L を最小化するように g を学習する。すなわち、近傍ユーザ群について、解釈用の線形回帰モデル g を、予測モデル f に近づくように学習する。ここで、 $\pi_x(x')$ は損失関数の重み係数を表しており、

[0058]

[数5]

$$\pi_x(x') = \exp\left(\frac{-D(x, x')^2}{\sigma^2}\right)$$

で定義される。距離が近いユーザほど損失関数を重視するため、より対象ユーザに近いユーザにフィッティングする解釈用モデルを学習する機構となる。解釈用モデルを学習した後、線形回帰モデル学習部41は、重要生活習慣返却部42へ解釈用モデルの偏回帰係数を引き渡す。なお、ある偏回帰係数は、その偏回帰係数に対応する項目以外の項目の値を固定してその偏回帰係数に対応する項目の値を増加させた場合の予測結果の増加量（もしくは減少量）を示す。

[0059] <S203：重要な生活習慣を返却>

S203において、重要生活習慣返却部42は、線形回帰モデル学習部41から渡された偏回帰係数に基づいて、重要な生活習慣を対象ユーザに返却（提示）する。具体的には下記のとおりである。

[0060] 生活習慣の一例として運動時間を i_1 とおく。このとき、 i_1 を変更したときに健康リスク値 g がどの程度変動するかを示す偏回帰係数

[0061] [数6]

$$\frac{\partial g}{\partial i_1}$$

の値が、運動時間の健康リスクに対する重要度となる。つまり、偏回帰係数の値が大きい生活習慣は、健康リスクへの影響が大きいということである。

[0062] 重要生活習慣返却部42は、線形回帰モデル学習部41が返却した偏回帰

係数のリスト

[0063] [数7]

$$\left(\frac{\partial g}{\partial i_1}, \frac{\partial g}{\partial i_2}, \dots, \frac{\partial g}{\partial K}\right)$$

を受け取り、受け取った偏回帰係数のリストに基づいて、ユーザ端末への情報提示を行う。

[0064] 例えば、重要生活習慣返却部42は、偏回帰係数及びそれに対応する生活習慣を偏回帰係数の降順で並び替え、ユーザインタフェース部300を介してユーザ端末へ返却する。ここで提示された生活習慣は、対象ユーザにとって改善すべき生活習慣となる。上にあるものほど重要なものである。

[0065] 図6、図7に、重要生活習慣返却部42からユーザ端末に提示される画面例を示す。

[0066] 図6は、偏回帰係数の降順で並び替えた生活習慣が表示された画面の例である。図6の例では、運動習慣（「1日30分以上の運動習慣がある」）が健康リスクに最も影響の大きな生活習慣であることがわかる。この例では、現状の生活習慣も表示されるようになっており、この対象ユーザは、「1日30分以上の運動習慣がある」は「はい」であることがわかる。

[0067] また、対象ユーザは、図6に示す画面上で、現状の生活習慣を変更することで、健康リスクの予測値がどの程度変動するかを確認する生活習慣の変更シミュレーションを行うことができる。

[0068] 図7は、ユーザ端末に表示される生活習慣の変更シミュレーション結果の画面例を示す。図7において、「before」で示される各健康リスク値は、現状の生活習慣から予測された値である。

[0069] 例えば、対象ユーザが、生活習慣 i_k を変更する。その変更量を Δi_k とする。変更の情報はユーザ端末から重要生活習慣返却部42に通知され、重要

生活習慣返却部 42 は、変更後の健康リスク値を、生活習慣 i_k に対応する偏回帰係数と Δi_k とを用いて

[0070] [数8]

$$(g + \frac{\partial g}{\partial k} \cdot \Delta i_k)$$

として計算し、ユーザインタフェース部 300 を介してユーザ端末の画面に表示する。図 7 の例では、生活習慣変更後の各健康リスク値が「after」として示されている。対象ユーザは、生活習慣の変更操作を繰り返すことで、生活習慣の改善と健康リスク値の変動との関係を視覚的に知ることができるようになる。

[0071] (重要生活習慣返却装置 100 の作用及び効果)

重要生活習慣返却装置 100 は、個々のユーザに対して健康リスクと関係の深い重要生活習慣を高い精度で提示することができる。

[0072] また、本実施の形態における重要生活習慣返却装置 100 は、結果の解釈と予測とを独立して行うことから、予測対象に合わせた予測モデルを開発者が選択することができる。

[0073] また、近傍ユーザを抽出する際に、単純に属性の近いユーザのみを抽出する場合には、妥当性の低い解釈結果が生成されることがあるが、本実施の形態にかかる重要生活習慣返却装置 100 では予測モデルを用いて近傍ユーザの絞り込みを行っているため、妥当性の高い解釈結果（精度の高い解釈結果）を得ることができる。

[0074] (他の適用分野の例)

以上、説明したように、本実施の形態における重要生活習慣返却装置 100 は、コントロール可能な変数であるユーザの生活習慣に基づいて、他の変数であるユーザの健康リスク値を予測するとともに、その予測結果を解釈す

ることで、生活習慣の改善効果を提示する。

[0075] このように、コントロール可能な変数に基づいて他の変数を予測する分野であれば、本発明により、健康以外の分野でも、予測、及び妥当性の高い予測結果の解釈を行うことが可能である。

[0076] 例えば、ユーザをターゲットとして動画を推薦するサービスにおいて、個々のユーザに対して、どのような動画推薦を行えばよいかをチューニングすることができる。この場合、動画の視聴履歴等をコントロール可能な変数として用い、推薦すべき動画を他の変数として予測し、本発明の手法により予測結果の解釈を行うことで、個々人に適した推薦方法をチューニングできる。

[0077] (実施の形態のまとめ)

以上、説明したように、本実施の形態により、複数ユーザのデータを格納するデータ記憶部と、前記複数ユーザ全体のデータから学習された予測モデルを格納するモデル記憶部と、対象ユーザに対する近傍ユーザを前記データ記憶部から抽出する近傍ユーザ探索部と、前記近傍ユーザを対象として、前記予測モデルに近似させた線形回帰モデルを学習する線形回帰モデル学習部と、前記線形回帰モデルの偏回帰係数に基づいて前記対象ユーザに対する予測の解釈結果を出力する解釈結果出力部と、を備え、前記近傍ユーザ探索部は、ユーザ間の距離に基づき抽出された近傍ユーザ候補を、前記予測モデルによる前記対象ユーザの予測方向に基づいて絞り込むことにより、前記近傍ユーザを抽出する予測解釈装置が提供される。

[0078] 上記の構成により、機械学習による予測と、予測の解釈とを高い精度で行うことが可能となる。なお、重要生活習慣返却装置100は、予測解釈装置の例である。重要生活習慣返却部42は、解釈結果出力部の例である。

[0079] 前記近傍ユーザ探索部は、前記予測方向として前記予測モデルの微分係数を算出し、前記近傍ユーザ候補についての状態と前記予測モデルによる予測結果とに対応する点と、前記対象ユーザについての状態と前記予測モデルによる予測結果とに対応する点とを結んだ直線の傾きを算出し、前記微分係数

と、前記直線の傾きとの積が負にならない近傍ユーザ候補を近傍ユーザとして抽出するようにしてもよい。この構成により、近傍ユーザを適切に抽出できる。

[0080] 前記解釈結果出力部は、前記線形回帰モデルの偏回帰係数の大きさの順位に基づいて、各偏回帰係数に対応する項目を並べて出力することができる。また、前記解釈結果出力部は、ある項目の値が変更された場合において、偏回帰係数に基づいて、変更後の予測結果を出力することができる。この構成により、例えば、ユーザは、項目の変更シミュレーションを行うことができる。

[0081] (実施形態の補足)

以上、本発明の実施の形態を説明してきたが、開示される発明はそのような実施形態に限定されず、当業者は様々な変形例、修正例、代替例、置換例等を理解するであろう。発明の理解を促すため具体的な数値例を用いて説明がなされたが、特に断りのない限り、それらの数値は単なる一例に過ぎず適切な如何なる値が使用されてもよい。上記の説明における項目の区分けは本発明に本質的ではなく、2以上の項目に記載された事項が必要に応じて組み合わせ使用されてよいし、ある項目に記載された事項が、別の項目に記載された事項に（矛盾しない限り）適用されてよい。機能ブロック図における機能部又は処理部の境界は必ずしも物理的な部品の境界に対応するとは限らない。複数の機能部の動作が物理的には1つの部品で行われてもよいし、あるいは1つの機能部の動作が物理的には複数の部品により行われてもよい。実施の形態で述べた処理手順については、矛盾の無い限り処理の順序を入れ替えてもよい。処理説明の便宜上、重要生活習慣返却装置100は機能的なブロック図を用いて説明されたが、そのような装置はハードウェアで、ソフトウェアで又はそれらの組み合わせで実現されてもよい。本発明の実施の形態に従ってユーザ装置10が有するプロセッサにより動作するソフトウェア及び本発明の実施の形態に従って基地局20が有するプロセッサにより動作するソフトウェアはそれぞれ、ランダムアクセスメモリ（RAM）、フラッ

シュメモリ、読み取り専用メモリ（ROM）、EPROM、EEPROM、レジスタ、ハードディスク（HDD）、リムーバブルディスク、CD-ROM、データベース、サーバその他の適切な如何なる記憶媒体に保存されてもよい。

[0082] 本明細書で使用する「に基づいて」という記載は、別段に明記されていない限り、「のみに基づいて」を意味しない。言い換えれば、「に基づいて」という記載は、「のみに基づいて」と「に少なくとも基づいて」の両方を意味する。

[0083] 「含む（include）」、「含んでいる（including）」、及びそれらの変形が、本明細書あるいは特許請求の範囲で使用されている限り、これら用語は、用語「備える（comprising）」と同様に、包括的であることが意図される。さらに、本明細書あるいは特許請求の範囲において使用されている用語「又は（or）」は、排他的論理和ではないことが意図される。

[0084] 本開示の全体において、例えば、英語での a、an 及び the のように、翻訳により冠詞が追加された場合、これらの冠詞は、文脈から明らかにそうではないことが示されていなければ、複数のものを含み得る。

[0085] 以上、本発明について詳細に説明したが、当業者にとっては、本発明が本明細書中に説明した実施形態に限定されるものではないということは明らかである。本発明は、特許請求の範囲の記載により定まる本発明の趣旨及び範囲を逸脱することなく修正及び変更態様として実施することができる。したがって、本明細書の記載は、例示説明を目的とするものであり、本発明に対して何ら制限的な意味を有するものではない。

[0086] 本特許出願は2018年8月24日に提出した日本国特許出願第2018-157718号に基づきその優先権を主張するものであり、日本国特許出願第2018-157718号の全内容を本願に援用する。

符号の説明

[0087] 100 重要生活習慣返却装置
200 予測モデル処理部

- 2 0 データ取得部
- 2 1 モデル予測部
- 2 2 データ記憶部
- 2 3 モデル記憶部
- 3 0 0 ユーザインタフェース部
- 4 0 0 解釈モデル処理部
- 4 0 近傍ユーザ探索部
- 4 1 線形回帰モデル学習部
- 4 2 重要生活習慣返却部
- 1 0 0 1 プロセッサ
- 1 0 0 2 メモリ
- 1 0 0 3 ストレージ
- 1 0 0 4 通信装置
- 1 0 0 5 入力装置
- 1 0 0 6 出力装置

請求の範囲

[請求項1]

複数ユーザのデータを格納するデータ記憶部と、
前記複数ユーザ全体のデータから学習された予測モデルを格納するモデル記憶部と、
対象ユーザに対する近傍ユーザを前記データ記憶部から抽出する近傍ユーザ探索部と、
前記近傍ユーザを対象として、前記予測モデルに近似させた線形回帰モデルを学習する線形回帰モデル学習部と、
前記線形回帰モデルの偏回帰係数に基づいて前記対象ユーザに対する予測の解釈結果を出力する解釈結果出力部と、を備え、
前記近傍ユーザ探索部は、ユーザ間の距離に基づき抽出された近傍ユーザ候補を、前記予測モデルによる前記対象ユーザの予測方向に基づいて絞り込むことにより、前記近傍ユーザを抽出する予測解釈装置。

[請求項2]

前記近傍ユーザ探索部は、
前記予測方向として前記予測モデルの微分係数を算出し、
前記近傍ユーザ候補についての状態と前記予測モデルによる予測結果とに対応する点と、前記対象ユーザについての状態と前記予測モデルによる予測結果とに対応する点とを結んだ直線の傾きを算出し、
前記微分係数と、前記直線の傾きとの積が負にならない近傍ユーザ候補を近傍ユーザとして抽出する
請求項1に記載の予測解釈装置。

[請求項3]

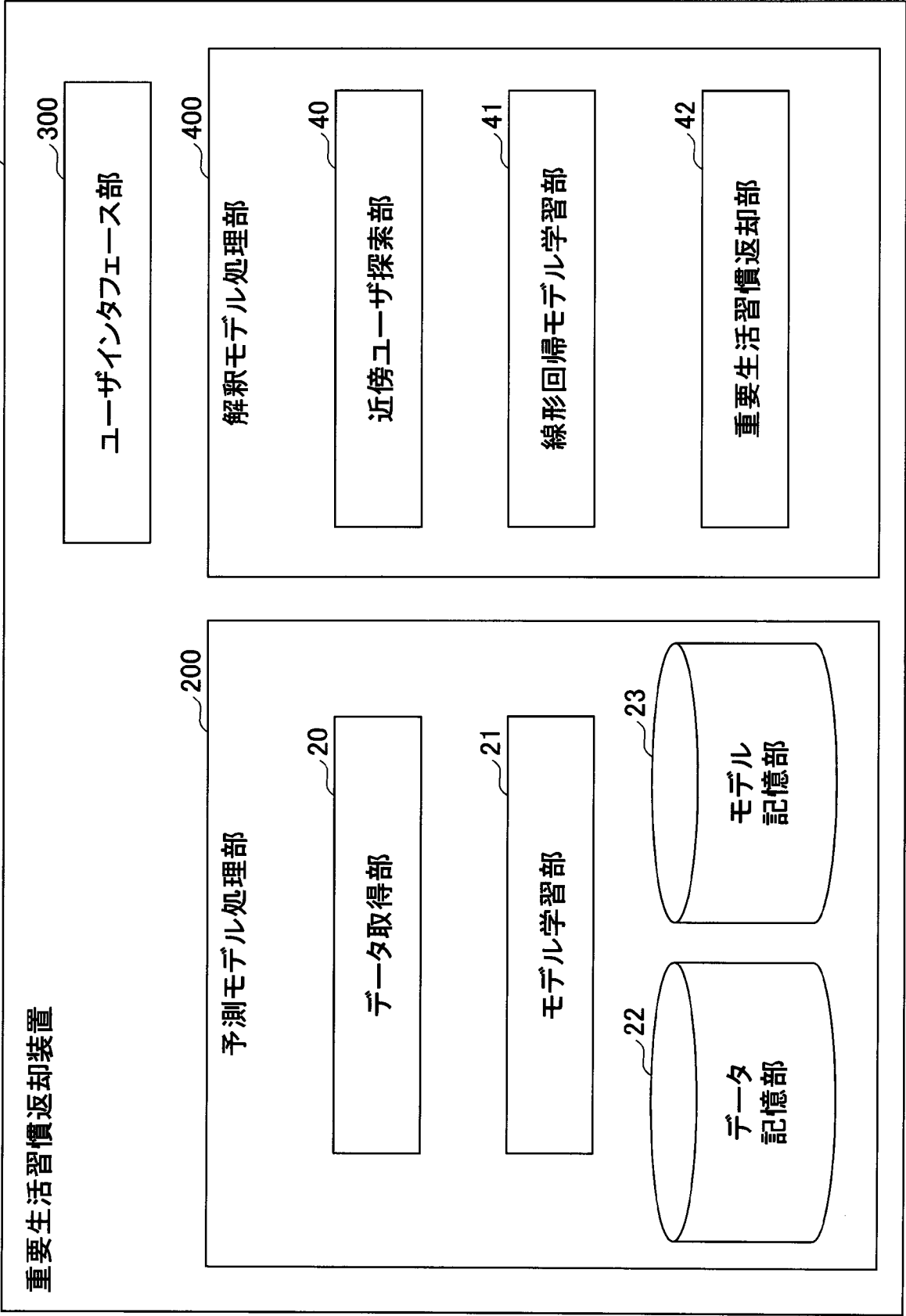
前記解釈結果出力部は、前記線形回帰モデルの偏回帰係数の大きさの順位に基づいて、各偏回帰係数に対応する項目を並べて出力する
請求項1又は2に記載の予測解釈装置。

[請求項4]

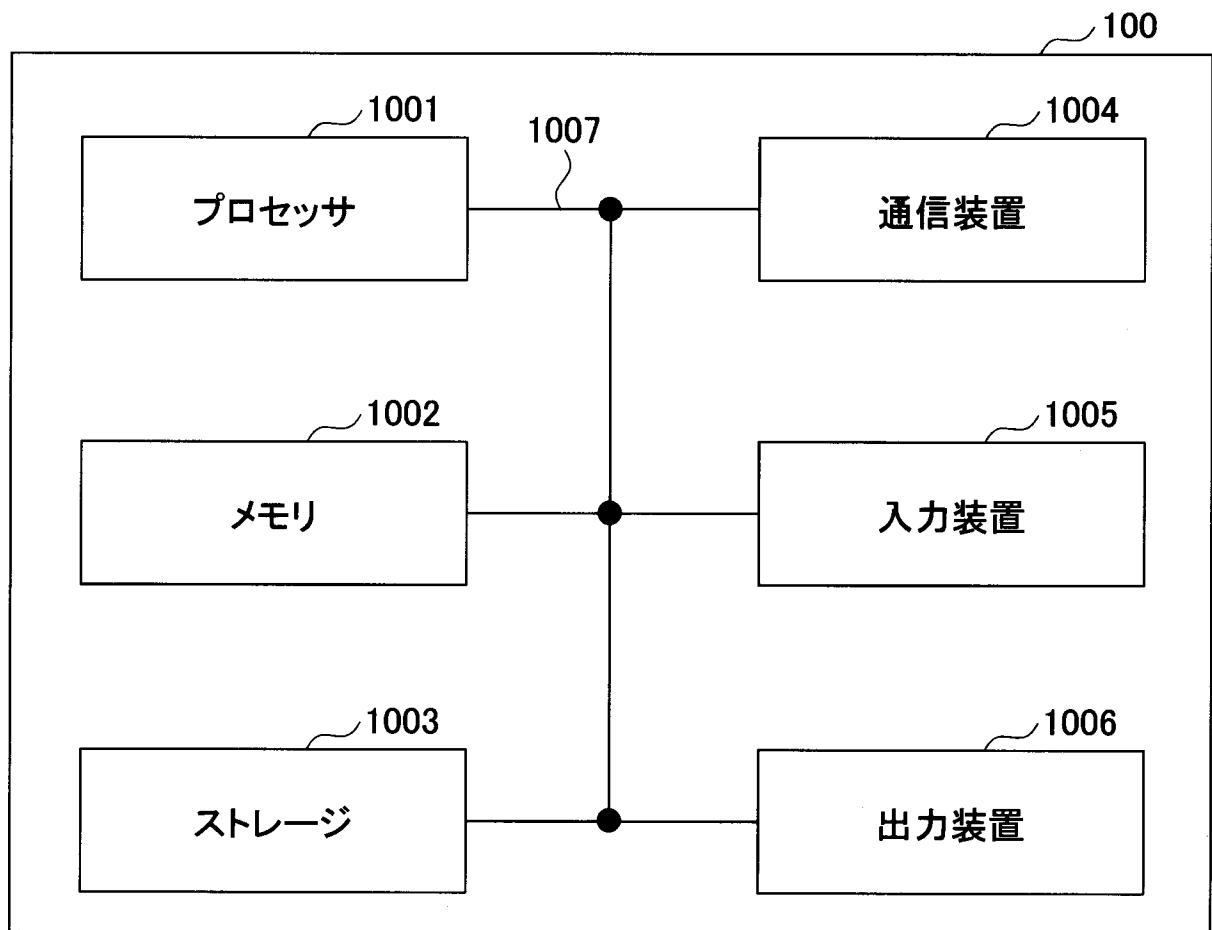
前記解釈結果出力部は、ある項目の値が変更された場合において、偏回帰係数に基づいて、変更後の予測結果を出力する
請求項3に記載の予測解釈装置。

- [請求項5] 複数ユーザのデータを格納するデータ記憶部と、前記複数ユーザ全体のデータから学習された予測モデルを格納するモデル記憶部とを備える予測解釈装置が実行する予測解釈方法であって、
- 対象ユーザに対する近傍ユーザを前記データ記憶部から抽出する探索ステップと、
- 前記近傍ユーザを対象として、前記予測モデルに近似させた線形回帰モデルを学習するステップと、
- 前記線形回帰モデルの偏回帰係数に基づいて前記対象ユーザに対する予測の解釈結果を出力するステップと、を備え、
- 前記探索ステップにおいて、前記予測解釈装置は、ユーザ間の距離に基づき抽出された近傍ユーザ候補を、前記予測モデルによる前記対象ユーザの予測方向に基づいて絞り込むことにより、前記近傍ユーザを抽出する
- 予測解釈方法。

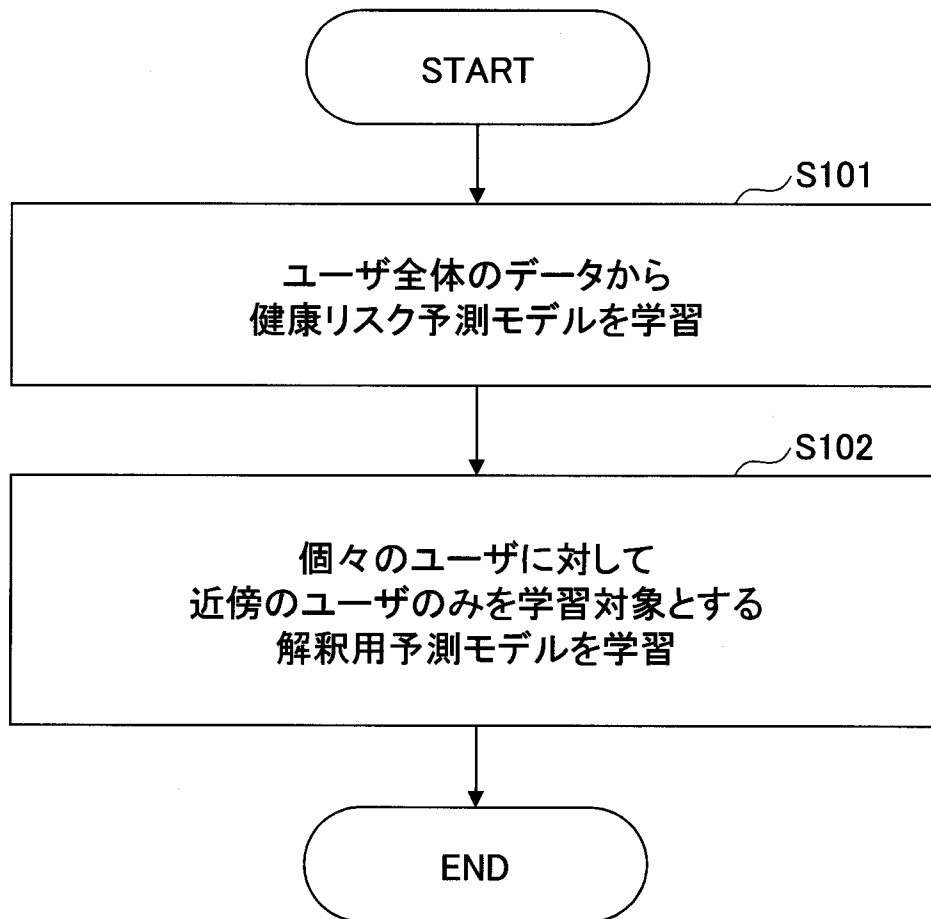
[図1]



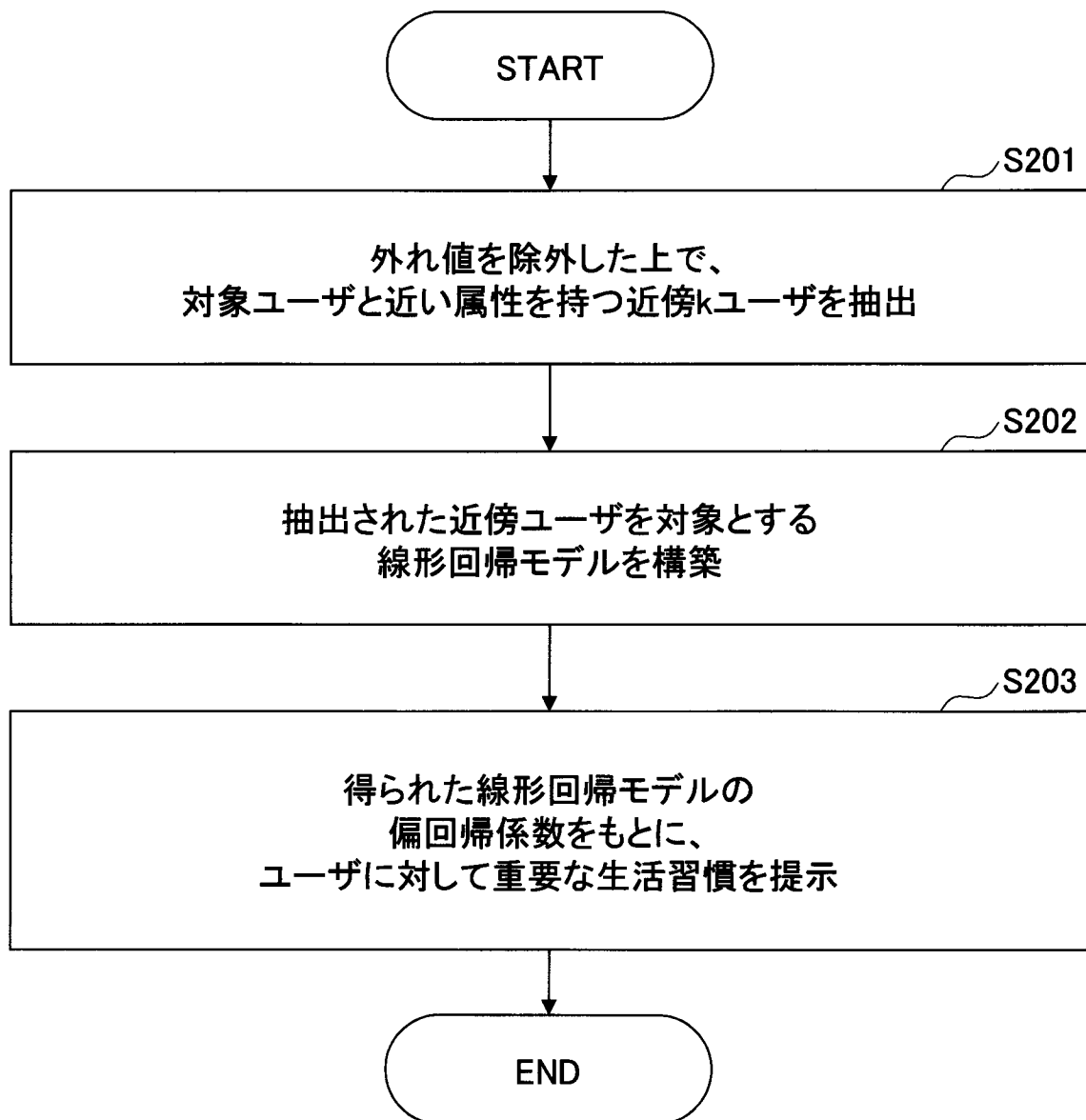
[図2]



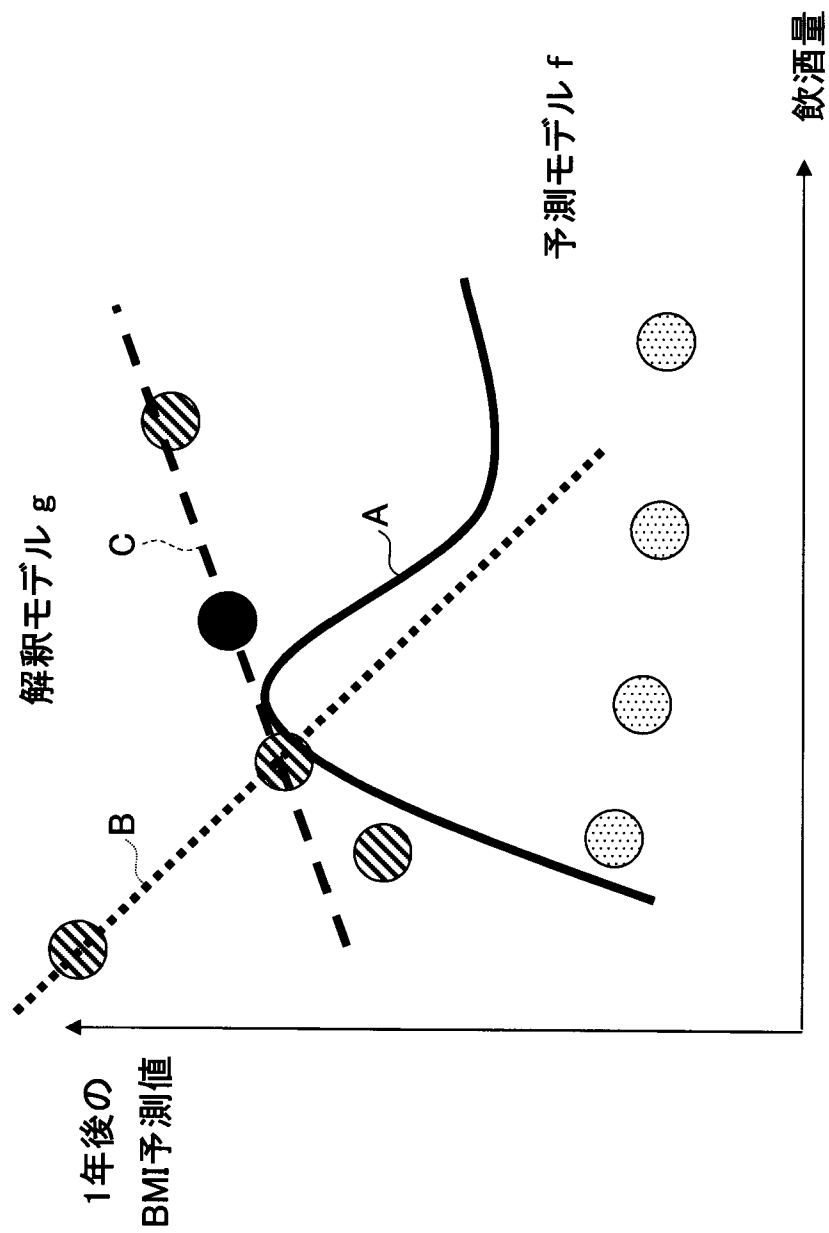
[図3]



[図4]



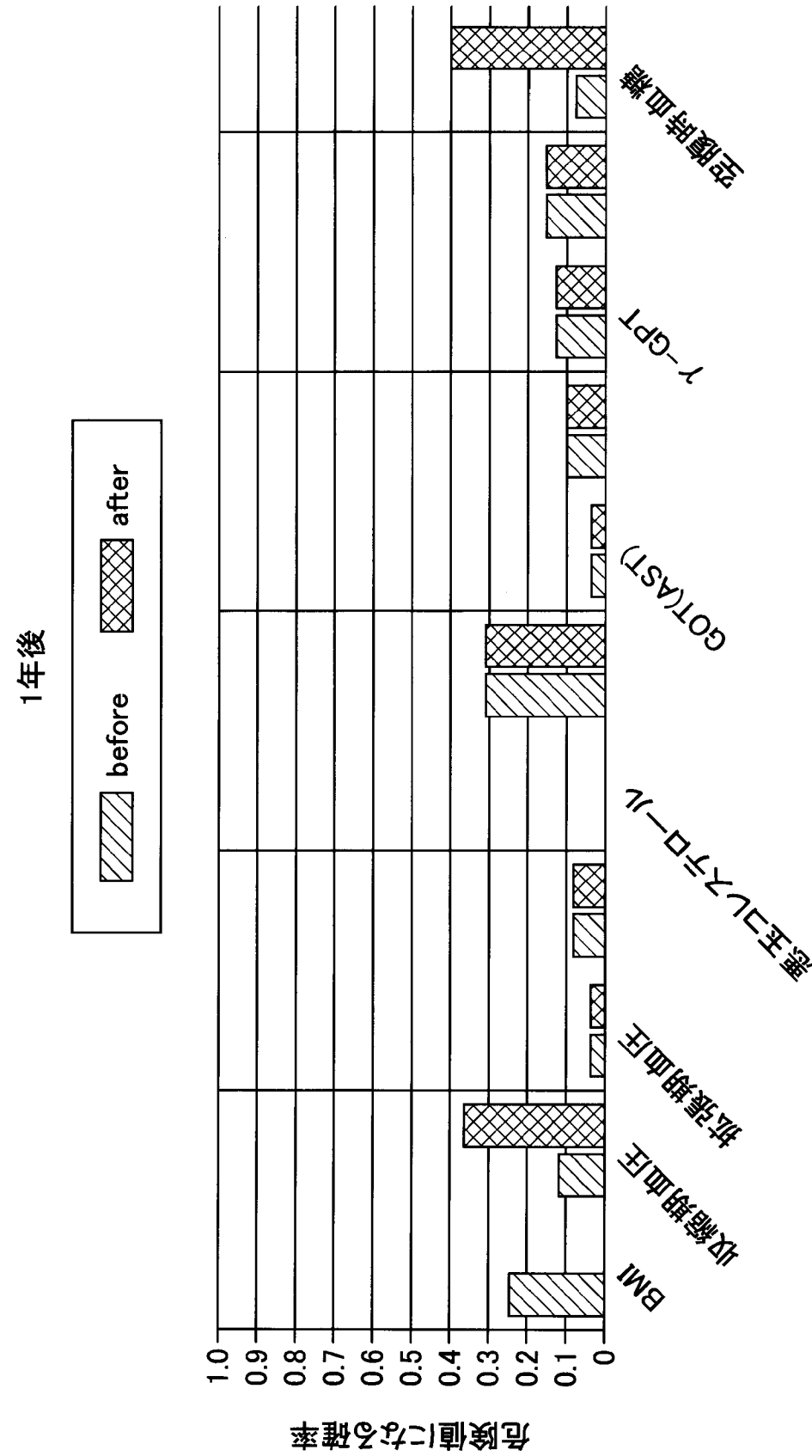
[図5]



[図6]

生活習慣	現状
1日30分以上の運動習慣がある	はい ✓
就寝2時間前以降の夕飯を 週3日以上行っている	はい ✓
習慣的にタバコを吸っている	いいえ ✓
20歳から10kg以上体重が増えている	いいえ ✓
歩行速度が、同性同年齢の人と 比較して速い	はい ✓
1年間の体重変化が±3kg以上ある	いいえ ✓
食べる早さ	ふつう ✓
夜食や間食を週3日以上行っている	いいえ ✓
飲酒習慣	時々 ✓
1日の飲酒量	1合～2合 ✓
睡眠で休養がとれる	はい ✓

[図7]



INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2019/032843

A. CLASSIFICATION OF SUBJECT MATTER

Int.Cl. G06Q10/04 (2012.01) i, G06N20/00 (2019.01) i, G16H50/30 (2018.01) i

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)

Int.Cl. G06Q10/00-99/00, G06N20/00, G16H10/00-80/00

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	1971-2019
Registered utility model specifications of Japan	1996-2019
Published registered utility model applications of Japan	1994-2019

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.
A	JP 2016-62509 A (YAHOO JAPAN CORPORATION) 25 April 2016, entire text, all drawings & US 2016/0086094 A1 entire text, all drawings	1-5
A	WO 2018/142753 A1 (NEC CORP.) 09 August 2018, entire text, all drawings (Family: none)	1-5
A	WO 2017/163277 A1 (NEC CORP.) 28 September 2017, entire text, all drawings & US 2019/0034945 A1 entire text, all drawings	1-5

☐ 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
11 November 2019 (11.11.2019)

Date of mailing of the international search report
19 November 2019 (19.11.2019)

Name and mailing address of the ISA/
Japan Patent Office
3-4-3, Kasumigaseki, Chiyoda-ku,
Tokyo 100-8915, Japan

Authorized officer

Telephone No.

A. 発明の属する分野の分類（国際特許分類（IPC））

Int.Cl. G06Q10/04(2012.01)i, G06N20/00(2019.01)i, G16H50/30(2018.01)i

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））

Int.Cl. G06Q10/00-99/00, G06N20/00, G16H10/00-80/00

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	1971-2019年
日本国実用新案登録公報	1996-2019年
日本国登録実用新案公報	1994-2019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 2016-62509 A（ヤフー株式会社）2016.04.25, 全文, 全図 & US 2016/0086094 A1 全文, 全図	1-5
A	WO 2018/142753 A1（日本電気株式会社）2018.08.09, 全文, 全図 （ファミリーなし）	1-5
A	WO 2017/163277 A1（日本電気株式会社）2017.09.28, 全文, 全図 & US 2019/0034945 A1 全文, 全図	1-5

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矛盾するものではなく、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

11.11.2019

国際調査報告の発送日

19.11.2019

国際調査機関の名称及びあて先

日本国特許庁（ISA/J P）

郵便番号100-8915

東京都千代田区霞が関三丁目4番3号

特許庁審査官（権限のある職員）

阿部 潤

5 R

4173

電話番号 03-3581-1101 内線 3502