

(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号
特許第7700565号
(P7700565)

(45)発行日 令和7年7月1日(2025.7.1)

(24)登録日 令和7年6月23日(2025.6.23)

(51)国際特許分類 F I
G 0 6 N 20/00 (2019.01) G 0 6 N 20/00
G 0 6 N 5/045(2023.01) G 0 6 N 5/045

請求項の数 6 (全14頁)

(21)出願番号	特願2021-129880(P2021-129880)	(73)特許権者	000005223 富士通株式会社 神奈川県川崎市中原区上小田中4丁目1番1号
(22)出願日	令和3年8月6日(2021.8.6)	(74)代理人	110002147 弁理士法人酒井国際特許事務所
(65)公開番号	特開2023-23919(P2023-23919A)	(72)発明者	徳安 慎也 福岡県福岡市早良区百道浜2丁目2番1号 富士通九州ネットワークテクノロジーズ株式会社内
(43)公開日	令和5年2月16日(2023.2.16)	(72)発明者	中島 哲 神奈川県川崎市中原区上小田中4丁目1番1号 富士通株式会社内
審査請求日	令和6年5月9日(2024.5.9)	審査官	新井 則和

最終頁に続く

(54)【発明の名称】 説明情報出力プログラム、説明情報出力方法および情報処理装置

(57)【特許請求の範囲】

【請求項1】

複数のデータのそれぞれを入力した場合の機械学習モデルの出力結果に対する、前記複数のデータのそれぞれに含まれる複数の因子それぞれの寄与度を取得し、

前記複数の因子のそれぞれの寄与度に基づいて、前記複数のデータをクラスタリングし、

前記クラスタリングによって生成された複数のグループのそれぞれについて、グループに含まれるデータを入力した場合の前記出力結果に対する前記複数の因子それぞれの寄与度の大きさを表す図を含む説明情報出力する、処理をコンピュータに実行させ、

前記出力する処理は、

前記複数のグループそれぞれについて、前記グループに含まれる各データの出力結果に対する前記複数の因子それぞれの寄与度の合計値を算出し、

前記複数のグループそれぞれについて、前記グループに含まれる前記複数の因子それぞれの寄与度の合計値に基づいた、前記複数の因子それぞれの寄与度の大きさを表す前記図を生成し、

前記複数のグループそれぞれに対応する前記図を含む前記説明情報出力することを特徴とする説明情報出力プログラム。

【請求項2】

前記出力する処理は、

前記複数のグループそれぞれについて、前記グループに含まれる前記複数の因子それぞれの寄与度の合計値を用いて、前記複数の因子それぞれが前記グループ内に占める割合を

10

20

算出し、

前記複数のグループそれぞれについて、前記複数の因子それぞれの割合を面積比で表した前記図を生成し、

前記複数のグループそれぞれに対応する前記図を含む前記説明情報出力プログラム、
ことを特徴とする請求項 1 に記載の説明情報出力プログラム。

【請求項 3】

前記出力する処理は、

前記複数の因子のうち指定された因子である複数の指定因子を軸とする特徴空間上で、前記複数の指定因子をベクトルとする因子ベクトルを用いて、前記複数のグループのうち第 1 のグループに含まれる各データを特定し、

前記各データに含まれる前記複数の因子それぞれの数を表した円グラフと前記第 1 のグループに対応する前記図とを含む前記説明情報出力プログラム、

ことを特徴とする請求項 2 に記載の説明情報出力プログラム。

【請求項 4】

前記出力する処理は、

前記複数のグループそれぞれについて、前記複数の因子のうち指定された因子である複数の指定因子を軸とする特徴空間上で、前記複数の指定因子をベクトルとする因子ベクトルを用いて各グループに含まれる各データを特定し、

前記複数のグループそれぞれについて、特定された各データに含まれる前記複数の因子それぞれの数を表した円グラフを生成し、

前記複数のグループそれぞれに対応する前記図と前記円グラフとを含む前記説明情報出力プログラム、

ことを特徴とする請求項 2 に記載の説明情報出力プログラム。

【請求項 5】

複数のデータのそれぞれを入力した場合の機械学習モデルの出力結果に対する、前記複数のデータのそれぞれに含まれる複数の因子それぞれの寄与度を取得し、

前記複数の因子のそれぞれの寄与度に基づいて、前記複数のデータをクラスタリングし、

前記クラスタリングによって生成された複数のグループのそれぞれについて、グループに含まれるデータを入力した場合の前記出力結果に対する前記複数の因子それぞれの寄与度の大きさを表す図を含む説明情報出力プログラム、処理をコンピュータが実行し、

前記出力する処理は、

前記複数のグループそれぞれについて、前記グループに含まれる各データの出力結果に対する前記複数の因子それぞれの寄与度の合計値を算出し、

前記複数のグループそれぞれについて、前記グループに含まれる前記複数の因子それぞれの寄与度の合計値に基づいた、前記複数の因子それぞれの寄与度の大きさを表す前記図を生成し、

前記複数のグループそれぞれに対応する前記図を含む前記説明情報出力プログラム、ことを特徴とする説明情報出力方法。

【請求項 6】

複数のデータのそれぞれを入力した場合の機械学習モデルの出力結果に対する、前記複数のデータのそれぞれに含まれる複数の因子それぞれの寄与度を取得し、

前記複数の因子のそれぞれの寄与度に基づいて、前記複数のデータをクラスタリングし、

前記クラスタリングによって生成された複数のグループのそれぞれについて、グループに含まれるデータを入力した場合の前記出力結果に対する前記複数の因子それぞれの寄与度の大きさを表す図を含む説明情報出力プログラム、

制御部を含み、

前記制御部は、

前記複数のグループそれぞれについて、前記グループに含まれる各データの出力結果に対する前記複数の因子それぞれの寄与度の合計値を算出し、

前記複数のグループそれぞれについて、前記グループに含まれる前記複数の因子それぞ

10

20

30

40

50

れの寄与度の合計値に基づいた、前記複数の因子それぞれの寄与度の大きさを表す前記図を生成し、

前記複数のグループそれぞれに対応する前記図を含む前記説明情報を出力することを特徴とする情報処理装置。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、説明情報出力プログラムなどに関する。

【背景技術】

【0002】

近年、機械学習（AI：Artificial Intelligence）により生成された機械学習モデルが利用されている。機械学習モデルは、仕組みの性質上、基本的に解釈が難しいという一面があり、それに対応するために説明可能なAI（XAI：Explainable AI）が利用されている。XAIは、機械学習モデルに対して入力した特徴量毎に因子寄与度を出し、どの特徴量によって予測結果や推定結果に至ったかを人間に対して説明可能に提示する技術である。

【先行技術文献】

【特許文献】

【0003】

【文献】特開2020-24542号公報

【文献】特開2020-135066号公報

【文献】米国特許出願公開第2021/27191号明細書

【文献】米国特許出願公開第2018/322955号明細書

【発明の概要】

【発明が解決しようとする課題】

【0004】

しかしながら、上記技術では、予測結果に対する説明情報として因子寄与度を算出することはできるが、全体の傾向を示すことは困難である。

【0005】

例えば、XAIは、AIの予測結果（インスタンス）ごとに、因子寄与度を出力するので、ユーザは、全体の傾向を把握するために、各予測結果と因子寄与度との関係を個別に確認することになる。この結果、ユーザが因子寄与度をもとに予測結果に対する対策を行う場合に、時間がかかり、予測結果に対する最適な対策も立て辛い。

【0006】

一つの側面では、機械学習モデルの出力結果に対する全体の傾向を示すことができる説明情報出力プログラム、説明情報出力方法および情報処理装置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】

【0007】

第1の案では、説明情報出力プログラムは、複数のデータのそれぞれを入力した場合の機械学習モデルの出力結果に対する、前記複数のデータのそれぞれに含まれる複数の因子それぞれの寄与度を取得し、前記複数の因子のそれぞれの寄与度に基づいて、前記複数のデータをクラスタリングし、前記クラスタリングによって生成された複数のグループのそれぞれについて、グループに含まれるデータを入力した場合の前記出力結果に対する前記複数の因子それぞれの寄与度の大きさを表す図を含む説明情報を出力する、処理をコンピュータに実行させることを特徴とする。

【発明の効果】

【0008】

一実施形態によれば、機械学習モデルの出力結果に対する全体の傾向を示すことができる。

10

20

30

40

50

【図面の簡単な説明】

【0009】

【図1】図1は、実施例1にかかる情報処理装置を説明する図である。

【図2】図2は、実施例1にかかる情報処理装置の機能構成を示す機能ブロック図である。

【図3】図3は、訓練データDBに記憶される情報の例を示す図である。

【図4】図4は、入力データDBに記憶される情報の例を示す図である。

【図5】図5は、因子寄与度の取得を説明する図である。

【図6】図6は、因子の比重の算出を説明する図である。

【図7】図7は、説明情報の表示例を説明する図である。

【図8】図8は、説明情報の表示例を説明する図である。

10

【図9】図9は、実施例1にかかる処理の流れを示すフローチャートである。

【図10】図10は、ハードウェア構成例を説明する図である。

【発明を実施するための形態】

【0010】

以下に、本願の開示する説明情報出力プログラム、説明情報出力方法および情報処理装置の実施例を図面に基づいて詳細に説明する。なお、この実施例によりこの発明が限定されるものではない。また、各実施例は、矛盾のない範囲内で適宜組み合わせることができる。

【0011】

図1は、実施例1にかかる情報処理装置10を説明する図である。図1に示す情報処理装置10は、LIMEなどのアルゴリズムが適用されたXAIを用いて、機械学習モデルの予測結果を、ユーザが視覚的に理解することができる説明情報に変換して出力するコンピュータの一例である。

20

【0012】

具体的には、情報処理装置10は、複数のデータのそれぞれを入力した場合の機械学習モデルの出力結果に対する、複数のデータのそれぞれに含まれる複数の因子それぞれの寄与度を取得する。情報処理装置10は、複数の因子のそれぞれの寄与度に基づいて、複数のデータをクラスタリングする。情報処理装置10は、クラスタリングによって生成された複数のグループのそれぞれについて、グループに含まれるデータを入力した場合の出力結果に対する複数の因子それぞれの寄与度の大きさを表す図を含む説明情報を出力する。

30

【0013】

例えば、図1に示すように、情報処理装置10は、特徴量A、特徴量B、特徴量C、特徴量Dを有する各入力データを機械学習モデルに入力して各予測結果を取得する。そして、情報処理装置10は、入力データ、予測結果およびXAIを用いて、各入力データに含まれる因子A、因子B、因子C、因子Dの各因子寄与度を取得する。ここで、因子寄与度は、各特徴量が予測結果に寄与した度合いを示す情報であり、因子Aは特徴量Aの寄与度を示し、因子Bは特徴量Bの寄与度を示し、因子Cは特徴量Cの寄与度を示し、因子Dは特徴量Dの寄与度を示す。

【0014】

続いて、情報処理装置10は、各入力データに対応する因子寄与度をクラスタリングする。例えば、情報処理装置10は、特徴量a、特徴量b、特徴量c、特徴量dを各次元(4次元)とする特徴空間で、因子A、因子B、因子C、因子Dの各ベクトルにより各入力データを特定し、各入力データをクラスタリングする。

40

【0015】

その後、情報処理装置10は、各クラスタについて、クラスタの全体に対する各因子の占める割合を面積比でソートして表示する。このように、情報処理装置10は、予測結果を機械学習モデルの出力の理由(因子ベクトル)でクラスタリングし、かつ、視覚的に分かり易い面積比などの形式で表示するので、機械学習モデルの出力結果に対する全体の傾向を示すことができる。

【0016】

50

図2は、実施例1にかかる情報処理装置10の機能構成を示す機能ブロック図である。図2に示すように、情報処理装置10は、通信部11、出力部12、記憶部13、制御部20を有する。

【0017】

通信部11は、他の装置との間の通信を制御する。例えば、通信部11は、管理者端末などから処理開始の指示や判定対象のデータ(入力データ)を受信し、制御部20による処理結果を管理者端末に送信する。

【0018】

出力部12は、各種情報を表示出力する。例えば、出力部12は、後述する機械学習モデル15の出力結果、制御部20により生成された説明情報などを表示出力する。

10

【0019】

記憶部13は、各種データや制御部20が実行するプログラムなどを記憶する。この記憶部13は、訓練データDB14、機械学習モデル15、入力データDB16を記憶する。

【0020】

訓練データDB14は、機械学習モデル15の機械学習に利用される訓練データを記憶するデータベースである。具体的には、訓練データDB14は、複数の特徴量と正解情報(ラベル)とを有する訓練データの集合を記憶する。一例として、通信事業者において契約情報から解約の可能性を判定する機械学習モデルの生成に利用される訓練データについて説明する。

【0021】

20

図3は、訓練データDB14に記憶される情報の例を示す図である。図3に示すように、訓練データDB14に記憶される各訓練データは、「会員ID、性別、年齢、契約期間、月額、年収、平均通信量、ラベル」を有する。ここで、「性別、年齢、契約期間、月額、年収、平均通信量」のそれぞれが特徴量であり、「ラベル」が正解情報である。なお、「会員ID」には、会員の識別子が設定され、「性別」には、会員の性別が設定され、「年齢」には、会員の年齢が設定され、「契約期間」には、会員の契約期間が設定される。「月額」には、会員が契約している月額料金が設定され、「年収」には、会員の年収が設定され、「平均通信量」には、会員が1か月に使用するデータ通信量の平均値が設定され、「ラベル」には、解約の有無が設定される。

【0022】

30

図3の例では、会員ID=1の会員には、「男性、40代、契約は2年以上、月額8000円、年収が800万円、平均5GBの通信量を使用」が設定されており、この会員が解約せずに「継続」していることが設定されている。

【0023】

機械学習モデル15は、複数の特徴量を有する入力されたデータに応じて判定結果を出力するように、訓練データDB14に記憶される訓練データを用いて生成される機械学習モデルである。上記通信事業者の例で説明すると、機械学習モデル15は、入力データが入力されると、解約する確率と解約しない確率とを出力する。なお、機械学習モデル15には、ニューラルネットワークやディープラーニングなどを採用することができる。

【0024】

40

入力データDB16は、機械学習モデル15に入力するデータであって、判定対象の入力データを記憶するデータベースである。上記通信事業者の例で説明すると、入力データDB16に記憶される各入力データは、解約するか否かを判定する対象の会員の特徴量を有するデータである。

【0025】

図4は、入力データDB16に記憶される情報の例を示す図である。図4に示すように、入力データDB16に記憶される各入力データは、「会員ID、性別、年齢、契約期間、月額、年収、平均通信量」を有する。ここで、「性別、年齢、契約期間、月額、年収、平均通信量」のそれぞれが特徴量である。なお、各特徴量の説明は、図3と同様なので、詳細な説明は省略する。図4の例では、会員ID=01の会員のデータには、特徴量とし

50

て、「男性、50代、5年以上、月額8000円、年収が1200万円、平均2GB」が設定されている。

【0026】

制御部20は、情報処理装置10全体を司る処理部であり、機械学習部21、予測部22、説明実行部23、表示制御部24を有する。

【0027】

機械学習部21は、訓練データDB14に記憶される各訓練データを用いて、機械学習モデル15を生成する。具体的には、機械学習部21は、訓練データを用いて教師あり学習により、機械学習モデル15を訓練する。図3の例で説明すると、機械学習部21は、訓練データDB14から会員IDの訓練データを取得し、性別などの特徴量を機械学習モデル15に入力する。そして、機械学習部21は、機械学習モデル15の出力値と、ラベル「継続（解約なし）」との誤差が小さくなるように、機械学習モデル15の機械学習を実行する。

10

【0028】

予測部22は、入力データDB16に記憶される各入力データに対して、機械学習モデル15を用いた予測を実行する。上記例で説明すると、予測部22は、入力データDB16から会員IDが01の入力データを取得し、性別などの特徴量を機械学習モデル15に入力する。そして、予測部22は、機械学習モデル15の出力結果を用いて、会員ID=01の会員が解約するか否かを予測する。なお、予測部22は、予測結果を出力部12に表示し、記憶部13に格納する。

20

【0029】

説明実行部23は、予測部22による各予測結果に対して、ユーザが確認可能な説明情報を生成する。具体的には、説明実行部23は、複数の入力データのそれぞれを入力した場合の機械学習モデル15の出力結果に対する、複数の入力データのそれぞれに含まれる複数の因子それぞれの寄与度を取得する。そして、説明実行部23は、複数の因子のそれぞれの寄与度に基づいて、複数の入力データをクラスタリングする。その後、説明実行部23は、クラスタリングによって生成された複数のグループのそれぞれについて、グループに含まれる入力データを入力した場合の出力結果に対する複数の因子それぞれの寄与度の大きさ（比重）を表す図を含む説明情報を生成する。

【0030】

30

まず、説明実行部23は、各入力データおよび予測結果と、XAIとを用いて、因子寄与度を取得する。図5は、因子寄与度の取得を説明する図である。図5に示すように、説明実行部23は、特徴量a、特徴量b、特徴量c、特徴量dを有する入力データを機械学習モデル15に入力し、予測結果を取得する。そして、説明実行部23は、入力データの特徴量を様々に変更させた各近傍データを生成して、各近傍データを機械学習モデル15に入力して各予測結果を取得する。

【0031】

続いて、説明実行部23は、入力データと予測結果、各近傍データと各予測結果をXAIに入力し、複雑な機械学習モデル15に対して入力データおよび近傍データを用いて局所的に近似する説明可能なモデル（線形回帰モデル）を生成する。そして、説明実行部23は、線形回帰モデルの偏回帰係数を計算することにより、特徴量aに対応する因子A、特徴量bに対応する因子B、特徴量cに対応する因子C、特徴量dに対応する因子Dの寄与度を取得する。

40

【0032】

このようにして、説明実行部23は、N個の入力データそれぞれに対して、予測結果と因子寄与度を取得する。なお、XAIを用いた因子寄与度の取得は、上記した処理に限定されるものではなく、LIMEなどのアルゴリズムなど公知の手法を採用することができる。

【0033】

次に、説明実行部23は、因子寄与度を用いて入力データをクラスタリングし、各クラ

50

スタについてそのクラスタ内の因子の比重を算出する。図6は、因子の比重の算出を説明する図である。図6に示すように、説明実行部23は、特徴量a、特徴量b、特徴量c、特徴量dを各軸とする4次元の特徴空間に、N個の入力データそれぞれに対するN個の因子寄与群をマッピングする。すなわち、説明実行部23は、因子Aの寄与度、因子Bの寄与度、因子Cの寄与度、因子Dの寄与度を各ベクトルとする因子ベクトルにより特定される特徴空間の位置に、各入力データをマッピングする。

【0034】

その後、説明実行部23は、特徴空間で入力データをクラスタリングし、クラスタ1、クラスタ2、クラスタ3などを生成する。そして、説明実行部23は、各クラスタ内における因子の比重を生成する。例えば、説明実行部23は、第1のグループの一例であるクラスタ4に属する各入力データの因子寄与度を取得し、因子Aの寄与度の合計、因子Bの寄与度の合計、因子Cの寄与度の合計、因子Dの寄与度の合計を算出する。そして、説明実行部23は、クラスタ4において各因子が占める因子Aの割合、因子Bの割合、因子Cの割合、因子Dの割合を算出する。

10

【0035】

このようにして、説明実行部23は、各クラスタにおける各因子が占める割合（比重）を面積比で表し、各面積比をソートした図を含む説明情報を生成して、表示制御部24に出力し、記憶部13に格納し、出力部12に出力する。

【0036】

表示制御部24は、説明実行部23により生成された説明情報を可視化して出力部12に表示出力する。例えば、表示制御部24は、特徴量を軸にしてクラスタ内のインスタンスをマッピングすることで、クラスタを細分化した情報を表示出力する。

20

【0037】

図7は、説明情報の表示例を説明する図である。図7に示すように、表示制御部24は、説明実行部23により生成されたクラスタ4内の因子の比重を示す図（図7の(a)）と、クラスタ4を細分化した図（図7の(b)）とをあわせて表示する。ここで、図7の(a)は、図6を用いて説明した手法により生成される因子の比重を示す図である。図7の(b)は、表示制御部24が生成する図である。例えば、表示制御部24は、クラスタ4内の各因子のうち、ユーザにより指定された指定因子である因子A（年代）と因子B（年収）とを軸にした2次元空間に、クラスタ4内の各入力データの因子Aと因子Bをベクトル（因子ベクトル）に用いて、各入力データをマッピングする。そして、表示制御部24は、各入力データに含まれる複数の因子それぞれの数を表した円グラフを生成する。

30

【0038】

より詳細には、表示制御部24は、クラスタ4内の入力データのうち、年齢が閾値以上である高齢者に対応する入力データについて、各因子の因子寄与度の合計値を算出し、その各因子の因子寄与度の合計値を用いた因子の割合を円グラフで生成する。同様に、表示制御部24は、年齢が閾値未満である若年層に対応する入力データについて、各因子の因子寄与度の合計値を算出し、その各因子の因子寄与度の合計値を用いた因子の割合を円グラフで生成して表示する。円グラフでは、円の面積でクラスタ内のインスタンス数が表現される。

40

【0039】

この結果、表示制御部24は、因子寄与度からクラスタリングしたクラスタに対して、特徴量を軸にしてクラスタ内のインスタンスをマッピングすることで、クラスタを細分化した説明情報を提示することができる。例えば、クラスタは因子寄与度でクラスタリングされているため、因子寄与度が近くても実際の特徴量が異なる場合がある。退会者予測を例に、因子寄与度の高い因子Aが年代だった場合を考えると、若年層または高齢者で、平均年収などに差があるが月額料金などの因子寄与度には差がないことがある。これに対して、表示制御部24がクラスタを細分化した情報を生成して表示することで、因子寄与度が高くない特徴量の因子寄与度の差分が明確になり、ユーザがその差分を視覚的に確認できる。

50

【 0 0 4 0 】

つまり、同一のクラスタ内には、同様の因子寄与度をもつユーザがクラスタリングされるが、実際の特徴量には差分が発生する可能性があるので、図7の(a)からのみでは読み取れないことも考えられる。また、クラスタ数の選択が適切でない場合、ユーザAの因子A(年代)が「因子寄与度 = 0.5、特徴量の実値 = 60代」、ユーザBの因子A(年代)が「因子寄与度 = 0.5、特徴量の実値 = 20代」のように、同一クラスタ内でも因子寄与度の傾向が異なる場合がある。

【 0 0 4 1 】

この場合であっても、表示制御部24により生成される図7の(b)に示すように、年代によってクラスタが二つに分割されたとすると、若年層と高齢者で平均年収や家族構成に差がある場合に、因子寄与度が下位の特徴量に微妙に傾向の差分が可視化され、ユーザはクラスタを細分化して確認することができる。また、ユーザは、図7の(a)のグラフだけでは、確認できない因子A(年代)の特徴量の実数値の傾向を確認することもできる。

10

【 0 0 4 2 】

別例としては、表示制御部24は、各クラスタを特徴量の軸にあわせてマッピングすることで、全体傾向の特徴をより具体的に可視化する。

【 0 0 4 3 】

図8は、説明情報の表示例を説明する図である。図8に示すように、表示制御部24は、説明実行部23により生成されたクラスタ4内の因子の比重を示す図(図8の(a))と、各クラスタのマッピング結果(図8の(b))とをあわせて表示する。ここで、図8の(a)は、図6を用いて説明した手法により生成される因子の比重を示す図である。図8の(b)は、表示制御部24が生成する図である。

20

【 0 0 4 4 】

例えば、表示制御部24は、各クラスタについて、因子の比重を円グラフで生成する。そして、表示制御部24は、各クラスタに対応する円グラフを、特徴量Aと特徴量Bの2次元空間にマッピングする。詳細に説明すると、表示制御部24は、クラスタ1内の入力データのうち、特徴量Aおよび特徴量Bが閾値以上である入力データについて、各因子の因子寄与度の合計値を算出し、その各因子の因子寄与度の合計値を用いた因子の割合を円グラフで生成する。同様に、表示制御部24は、クラスタ2内の入力データのうち、特徴量Aが閾値以上および特徴量Bが閾値未満である入力データについて、各因子の因子寄与度の合計値を算出し、その各因子の因子寄与度の合計値を用いた因子の割合を円グラフで生成する。

30

【 0 0 4 5 】

この結果、ユーザは、例えば退会予測なら契約期間、年代、性別など、対策を行いたい特徴量を多く持つクラスタを確認でき、対策が立てやすくなる。軸にする特徴量については、因子寄与度の高い特徴量が採用されてもよく、あるいは、ユーザが任意に選択できる。

【 0 0 4 6 】

例えば、図8の(a)に示すグラフでは、各クラスタの因子寄与度の傾向を確認することができるが、各因子の特徴量の実数値や、各クラスタに何人のユーザが含まれているかを視覚的に確認する事が難しい。これに対して、表示制御部24により生成される図8の(b)に示す円グラフにより、クラスタ4に年代が高く契約期間が長いユーザが多い、クラスタ4はユーザが多く、クラスタ1はユーザが少ないなど、ユーザは視覚的に容易に確認できる。

40

【 0 0 4 7 】

図9は、実施例1にかかる処理の流れを示すフローチャートである。図9に示すように、情報処理装置10の制御部20は、処理開始が指示されると(S101: Yes)、訓練データを用いて機械学習モデル15を生成する(S102)。

【 0 0 4 8 】

続いて、情報処理装置10の制御部20は、入力データを機械学習モデル15に入力し(S103)、予測結果を取得し(S104)、XAIなどを用いて因子寄与度を取得す

50

る (S 1 0 5)。ここで、未処理の入力データが存在する場合 (S 1 0 6 : Y e s)、制御部 2 0 は、S 1 0 3 に戻って、次の入力データについて以降の処理を実行する。

【 0 0 4 9 】

一方、未処理の入力データが存在しない場合 (S 1 0 6 : N o)、制御部 2 0 は、因子寄与度を用いて、入力データをクラスタリングする (S 1 0 7)。そして、制御部 2 0 は、クラスタリングされた各クラスタについて因子の比重を算出し (S 1 0 8)、説明情報を表示する説明画面を生成して出力する (S 1 0 9)。

【 0 0 5 0 】

上述したように、情報処理装置 1 0 は、予測結果 (インスタンス) をクラスタに分類し、それぞれのクラスタに対する因子寄与度の割合で出力することができる。この結果、ユーザは、予測結果を確認する際に、表示された因子寄与度の順に従って確認できるので、全体の予測結果の傾向を把握することができる。

10

【 0 0 5 1 】

また、予測のインプットとなる特徴量でクラスタリングした場合、それぞれの特徴量の重みは平等になる。しかし、情報処理装置 1 0 は、因子寄与度のベクトルでクラスタリングするので、因子寄与度により各特徴量に重みづけを行うことができる。この結果、情報処理装置 1 0 は、クラスタ内の予測結果と因子寄与度の傾向とを対応付けて表示することができ、ユーザの視認性を向上させることができる。

【 0 0 5 2 】

上記実施例で用いたデータ例、クラスタ数、特徴量、特徴量の数、因子、グラフ例、画面例等は、あくまで一例であり、任意に変更することができる。なお、クラスタは、グループの一例である。また、因子寄与度の大きさの一例として、面積、比重などを例示したが、これに限定されるものではなく、例えば数値、クラスタ内の合計値、平均値などの各指標を用いることもできる。また、図 7 や図 8 で説明した特徴量の軸も任意に設定変更することができる。

20

【 0 0 5 3 】

また、上記実施例では、通信事業者の解約を例にして説明したが、これに限定されるものではない。例えば、情報処理装置 1 0 は、音声データや画像データを用いた不審者検知など、様々な分析に適用することができる。

【 0 0 5 4 】

上記文書中や図面中で示した処理手順、制御手順、具体的名称、各種のデータやパラメータを含む情報については、特記する場合を除いて任意に変更されてもよい。

30

【 0 0 5 5 】

また、各装置の構成要素の分散や統合の具体的な形態は図示のものに限られない。例えば、説明実行部 2 3 と表示制御部 2 4 とが統合されてもよい。つまり、その構成要素の全部または一部は、各種の負荷や使用状況などに応じて、任意の単位で機能的または物理的に分散・統合されてもよい。さらに、各装置の各処理機能は、その全部または任意の一部が、CPU および当該 CPU にて解析実行されるプログラムにて実現され、あるいは、ワイヤードロジックによるハードウェアとして実現され得る。

【 0 0 5 6 】

図 1 0 は、ハードウェア構成例を説明する図である。図 1 0 に示すように、情報処理装置 1 0 は、通信装置 1 0 a、HDD (Hard Disk Drive) 1 0 b、メモリ 1 0 c、プロセッサ 1 0 d を有する。また、図 1 0 に示した各部は、バス等で相互に接続される。なお、情報処理装置 1 0 は、ディスプレイやタッチパネルなどを有していてもよい。

40

【 0 0 5 7 】

通信装置 1 0 a は、ネットワークインタフェースカードなどであり、他の装置との通信を行う。HDD 1 0 b は、図 2 に示した機能を動作させるプログラムや DB を記憶する。

【 0 0 5 8 】

プロセッサ 1 0 d は、図 2 に示した各処理部と同様の処理を実行するプログラムを HDD 1 0 b 等から読み出してメモリ 1 0 c に展開することで、図 2 等で説明した各機能を実

50

行するプロセスを動作させる。例えば、このプロセスは、情報処理装置 10 が有する各処理部と同様の機能を実行する。具体的には、プロセッサ 10 d は、機械学習部 21、予測部 22、説明実行部 23、表示制御部 24 等と同様の機能を有するプログラムを HDD 10 b 等から読み出す。そして、プロセッサ 10 d は、機械学習部 21、予測部 22、説明実行部 23、表示制御部 24 等と同様の処理を実行するプロセスを実行する。

【0059】

このように、情報処理装置 10 は、プログラムを読み出して実行することで説明情報出力方法を実行する情報処理装置として動作する。また、情報処理装置 10 は、媒体読取装置によって記録媒体から上記プログラムを読み出し、読み出された上記プログラムを実行することで上記した実施例と同様の機能を実現することもできる。なお、この他の実施例

10

【0060】

このプログラムは、インターネットなどのネットワークを介して配布されてもよい。また、このプログラムは、ハードディスク、フレキシブルディスク (FD)、CD-ROM、MO (Magneto - Optical disk)、DVD (Digital Versatile Disc) などのコンピュータで読み取り可能な記録媒体に記録され、コンピュータによって記録媒体から読み出されることによって実行されてもよい。

【符号の説明】

20

【0061】

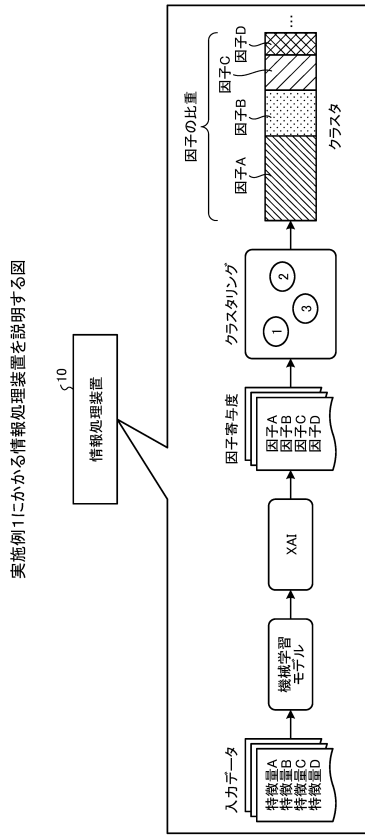
- 10 情報処理装置
- 11 通信部
- 12 出力部
- 13 記憶部
- 14 訓練データ DB
- 15 機械学習モデル
- 16 入力データ DB
- 20 制御部
- 21 機械学習部
- 22 予測部
- 23 説明実行部
- 24 表示制御部

30

40

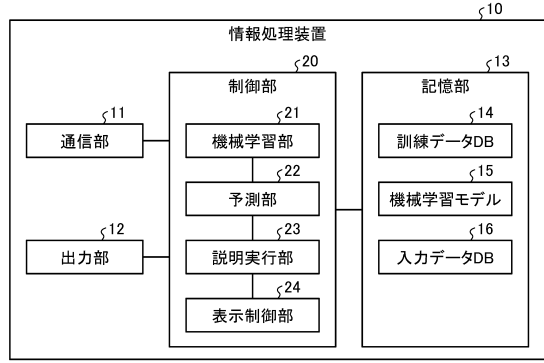
50

【図面】
【図 1】



【図 2】

実施例1にかかる情報処理装置の機能構成を示す機能ブロック図



10

20

【図 3】

訓練データDBに記憶される情報の例を示す図

会員ID	性別	年齢	契約期間	月額	年収	平均通信料	ラベル
1	男性	40代	2年以上	5000円	800万円	5GB	継続 (解約なし)
2	女性	30代	1年以上	5000円	700万	8GB	解約
...

【図 4】

入力データDBに記憶される情報の例を示す図

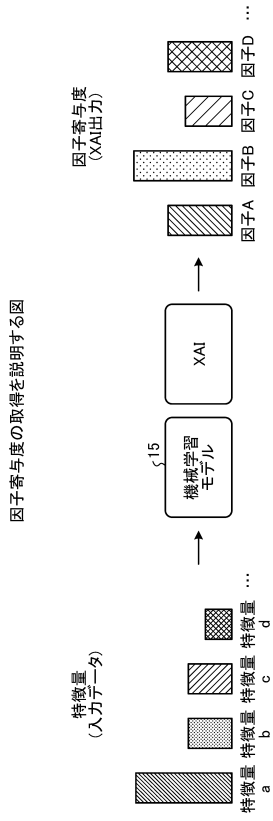
会員ID	性別	年齢	契約期間	月額	年収	平均通信料
01	男性	50代	5年以上	8000円	1200万円	2GB
...

30

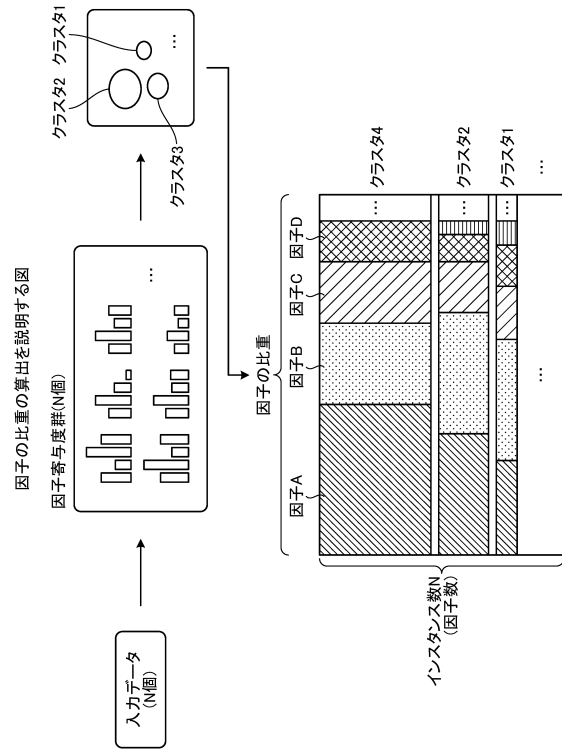
40

50

【 図 5 】

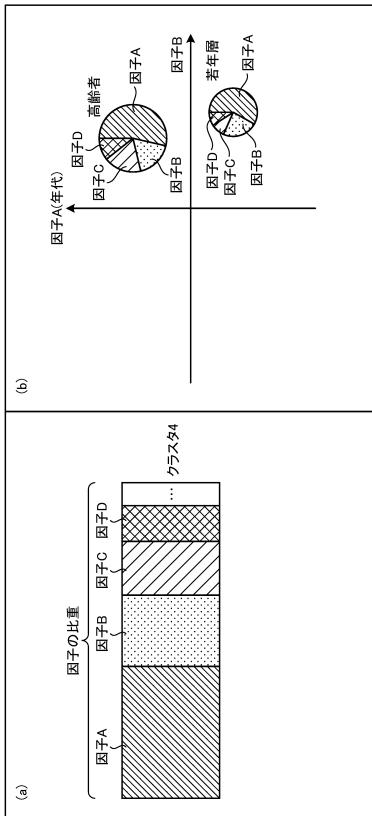


【 図 6 】



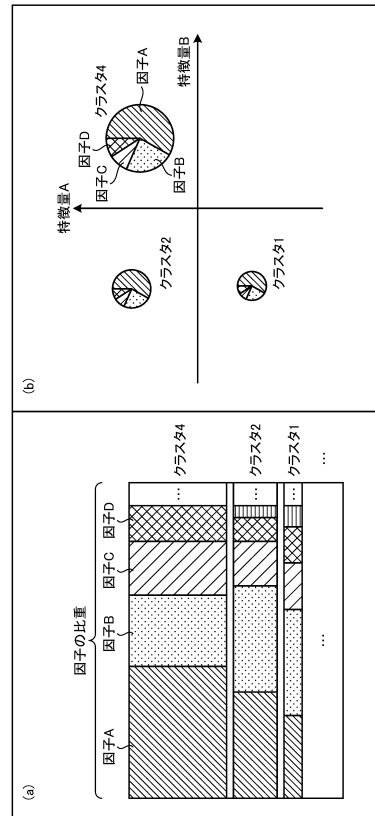
【 図 7 】

説明情報の表示例を説明する図



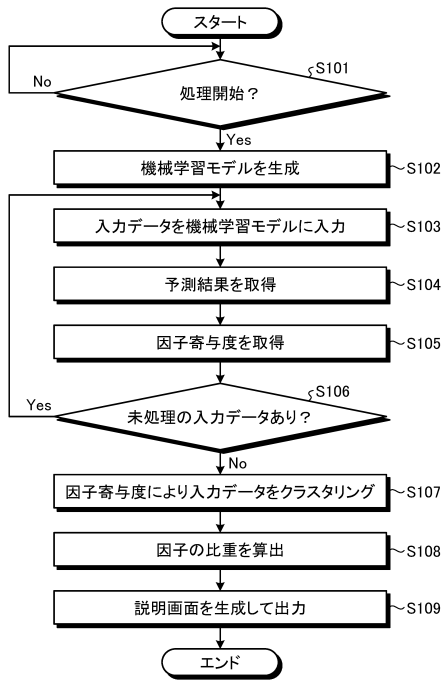
【 図 8 】

説明情報の表示例を説明する図



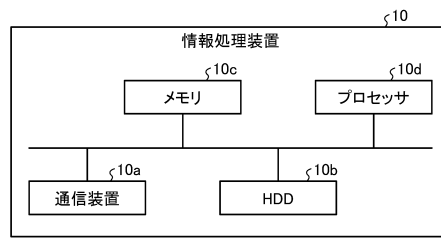
【 図 9 】

実施例1にかかる処理の流れを示すフローチャート



【 図 1 0 】

ハードウェア構成例を説明する図



10

20

30

40

50

フロントページの続き

- (56)参考文献 国際公開第 2 0 2 1 / 0 9 6 5 6 4 (W O , A 1)
GROMIT YEUK-YIN CHAN et al. , SUBPLEX: Towards a Better Understanding of BlackBoxM
odel Explanations at the Subpopulation Level , arxiv.org, [online] , 2020年07月21日 , pp.
111:1-111:28 , [検索日 2025.02.20], Retrieved from the Internet: URL: [https://arxiv.org
/pdf/2007.10609v1](https://arxiv.org/pdf/2007.10609v1)
西井 進剛 , 知識集約型企業のグローバル戦略とビジネスモデル , 2013年05月20日 , pp.
196-216
- (58)調査した分野 (Int.Cl. , D B 名)
G 0 6 N 2 0 / 0 0
G 0 6 N 5 / 0 4 5