

【公報種別】特許法第17条の2の規定による補正の掲載  
 【部門区分】第6部門第3区分  
 【発行日】令和6年9月26日(2024.9.26)

【国際公開番号】WO2024/096054  
 【出願番号】特願2024-520737(P2024-520737)  
 【国際特許分類】  
 G 0 6 Q 3 0 / 0 2 ( 2 0 2 3 . 0 1 )  
 【 F I 】  
 G 0 6 Q 3 0 / 0 2

10

【手続補正書】  
 【提出日】令和6年4月4日(2024.4.4)

【手続補正1】  
 【補正対象書類名】特許請求の範囲  
 【補正対象項目名】全文  
 【補正方法】変更  
 【補正の内容】  
 【特許請求の範囲】

【請求項1】

20

ユーザに関する複数の変数を有する第1のデータソースを取得する取得ステップであって、前記複数の変数には前記ユーザの属性情報に関する変数が含まれる、取得ステップと

、  
 前記第1のデータソースに含まれる前記複数の変数から説明変数および目的変数を選択する選択ステップであって、前記説明変数には、前記属性情報に関する変数と、少なくとも1つの前記属性情報とは異なる変数が含まれており、前記目的変数は複数選択される、選択ステップと、

前記第1のデータソースを教師データとして、選択された前記説明変数から前記複数の目的変数を予測する予測モデルを作成する、作成ステップと、

前記予測モデルを、選択された前記説明変数を変数として含む第2のデータソースに適用して、疑似シングルソースデータを生成する、生成ステップと、  
 を有することを特徴とする情報処理方法。

30

【請求項2】

前記疑似シングルソースデータは、同一の前記ユーザに関する複数の変数を有するシングルソースデータを擬似的に生成したものであることを特徴とする請求項1に記載の情報処理方法。

【請求項3】

前記生成ステップにおいては、前記ユーザの個人識別子を含まない前記第2のデータソースを用いて前記疑似シングルソースデータを生成することを特徴とする請求項1または2に記載の情報処理方法。

40

【請求項4】

前記作成ステップにおいては、選択された複数の前記説明変数と、選択された複数の前記目的変数と、の間の関係性を示す情報を利用して、前記予測モデルを作成し、

前記生成ステップにおいては、前記予測モデルを前記第2のデータソースに適用して、前記疑似シングルソースデータを生成することを特徴とする請求項1または2に記載の情報処理方法。

【請求項5】

前記作成ステップにおいては、選択された複数の前記説明変数と、選択された複数の前記目的変数と、の間の分散共分散構造を利用して、前記予測モデルを作成し、

前記生成ステップにおいては、前記予測モデルを前記第2のデータソースに適用して、

50

前記第 1 のデータソースの分散共分散構造に類似する分散共分散構造を持つ疑似シングルソースデータを生成する

ことを特徴とする請求項 4 に記載の情報処理方法。

【請求項 6】

前記作成ステップにおいては、選択された複数の前記変数の、平均値、歪度、および尖度のうち少なくとも 1 つをさらに用いて、前記予測モデルを作成する

ことを特徴とする請求項 5 に記載の情報処理方法。

【請求項 7】

前記作成ステップにおいては、前記選択ステップで前記第 1 のデータソースから選択された複数の前記変数のうち少なくとも 1 つの変数に対してノイズを付加したのち、前記予測モデルを作成する

ことを特徴とする請求項 1 または 2 に記載の情報処理方法。

10

【請求項 8】

前記作成ステップにおいては、ラプラスメカニズム、ガウシアンメカニズム、または指数メカニズムのいずれかを用いて、前記ノイズを付加する

ことを特徴とする請求項 7 に記載の情報処理方法。

【請求項 9】

前記作成ステップにおいては、前記ユーザに対して差分プライバシーの機能を提供するように、前記第 1 のデータソースに含まれる前記変数に対して前記ノイズを付加する

ことを特徴とする請求項 7 に記載の情報処理方法。

20

【請求項 10】

ユーザに関する複数の変数を有する第 1 のデータソースを取得する取得ステップであって、前記複数の変数には前記ユーザの属性情報に関する変数が含まれる、取得ステップと、

前記第 1 のデータソースに含まれる前記複数の変数から説明変数および目的変数を選択する選択ステップであって、前記説明変数には、前記属性情報に関する変数と、少なくとも 1 つの前記属性情報以外の変数とが含まれており、前記目的変数は複数選択される、選択ステップと、

前記第 1 のデータソースを元に、選択された前記複数の説明変数と前記複数の目的変数を含むデータの分散共分散行列を計算する、計算ステップと、

30

前記第 1 のデータソースとは異なる第 2 のデータソースであって、変数として、前記第 1 のデータソースで選択された前記複数の説明変数を含む第 2 のデータソースを取得する、第 2 の取得ステップと、

前記第 2 のデータソースに最適化計算を適用して前記複数の目的変数の値を生成し、生成された前記複数の目的変数の値を含む疑似シングルソースデータを生成する、生成ステップと、

を有することを特徴とする情報処理方法。

【請求項 11】

前記生成ステップにおいては、前記計算ステップで計算された前記分散共分散行列と、前記複数の変数の平均値が再現されるように、式 (1) および式 (2) を用いて最適化計算を行うものであり、

40

前記式 (3) および式 (4) に示す勾配法を用いて前記最適化計算を行うことを特徴とする請求項 10 に記載の情報処理方法。

【数 1】

$$\min_{Y_j} \lambda |\hat{\sigma}_{ij} - (X_i - \hat{\mu}_i^{(x)})^t (Y_j - \hat{\mu}_j^{(y)}) / (n-1)| + \kappa (Y_j^t \mathbf{1} / n - \hat{\mu}_j^{(y)})^2 + \nu \|Y_j \circ (Y_j - \mathbf{1})\|_2^2$$

… (1)

【数 2】

50

$$\min_{Y_j} \lambda |\hat{\sigma}_{jk} - (\mathbf{Y}_k - \hat{\boldsymbol{\mu}}_k^{(y)})^t (\mathbf{Y}_j - \hat{\boldsymbol{\mu}}_j^{(y)}) / (n-1)| + \kappa (\mathbf{Y}_j^t \mathbf{1} / n - \hat{\boldsymbol{\mu}}_j^{(y)})^2 + \nu \|\mathbf{Y}_j \circ (\mathbf{Y}_j - \mathbf{1})\|_2^2$$

… (2)

【数 3】

$$\frac{\partial f}{\partial \mathbf{Y}_j} = \begin{cases} \frac{-1}{n-1} (\mathbf{X}_i - \hat{\boldsymbol{\mu}}_i^{(x)}) & (|\hat{\sigma}_{ij} - (\mathbf{X}_i - \hat{\boldsymbol{\mu}}_i^{(x)})^t (\mathbf{Y}_j - \hat{\boldsymbol{\mu}}_j^{(y)}) / (n-1)| \geq 0) \\ \frac{1}{n-1} (\mathbf{X}_i - \hat{\boldsymbol{\mu}}_i^{(x)}) & (|\hat{\sigma}_{ij} - (\mathbf{X}_i - \hat{\boldsymbol{\mu}}_i^{(x)})^t (\mathbf{Y}_j - \hat{\boldsymbol{\mu}}_j^{(y)}) / (n-1)| < 0) \end{cases}$$

$$\frac{\partial g}{\partial \mathbf{Y}_j} = \frac{2}{n} (\mathbf{Y}_j^t \mathbf{1} / n - \hat{\boldsymbol{\mu}}_j^{(y)}) \mathbf{1}$$

$$\frac{\partial h}{\partial \mathbf{Y}_j} = 2 \mathbf{Y}_j \circ (\mathbf{Y}_j - \mathbf{1}) \circ (2 \mathbf{Y}_j - \mathbf{1})$$

$$\mathbf{Y}_j^{(n+1)} \leftarrow \mathbf{Y}_j^{(n)} - \lambda \frac{\partial f}{\partial \mathbf{Y}_j} - \kappa \frac{\partial g}{\partial \mathbf{Y}_j} - \nu \frac{\partial h}{\partial \mathbf{Y}_j}$$

… (3)

【数 4】

$$\frac{\partial f}{\partial \mathbf{Y}_j} = \begin{cases} \frac{-1}{n-1} (\mathbf{Y}_k - \hat{\boldsymbol{\mu}}_k^{(y)}) & (|\hat{\sigma}_{kj} - (\mathbf{Y}_k - \hat{\boldsymbol{\mu}}_k^{(y)})^t (\mathbf{Y}_j - \hat{\boldsymbol{\mu}}_j^{(y)}) / (n-1)| \geq 0) \\ \frac{1}{n-1} (\mathbf{Y}_k - \hat{\boldsymbol{\mu}}_k^{(y)}) & (|\hat{\sigma}_{kj} - (\mathbf{Y}_k - \hat{\boldsymbol{\mu}}_k^{(y)})^t (\mathbf{Y}_j - \hat{\boldsymbol{\mu}}_j^{(y)}) / (n-1)| < 0) \end{cases}$$

$$\frac{\partial g}{\partial \mathbf{Y}_j} = \frac{2}{n} (\mathbf{Y}_j^t \mathbf{1} / n - \hat{\boldsymbol{\mu}}_j^{(y)}) \mathbf{1}$$

$$\frac{\partial h}{\partial \mathbf{Y}_j} = 2 \mathbf{Y}_j \circ (\mathbf{Y}_j - \mathbf{1}) \circ (2 \mathbf{Y}_j - \mathbf{1})$$

$$\mathbf{Y}_j^{(n+1)} \leftarrow \mathbf{Y}_j^{(n)} - \lambda \frac{\partial f}{\partial \mathbf{Y}_j} - \kappa \frac{\partial g}{\partial \mathbf{Y}_j} - \nu \frac{\partial h}{\partial \mathbf{Y}_j}$$

… (4)

【請求項 1 2】

オブジェクトに関する複数の変数を有する第 1 のデータソースを取得する取得ステップであって、前記複数の変数には前記オブジェクトの属性情報に関する変数が含まれる、取得ステップと、

前記第 1 のデータソースに含まれる前記複数の変数から説明変数および目的変数を選択する選択ステップであって、前記説明変数には、前記属性情報に関する変数と、少なくとも 1 つの前記属性情報とは異なる変数が含まれており、前記目的変数は複数選択される、選択ステップと、

前記第 1 のデータソースを教師データとして、選択された前記説明変数から前記複数の目的変数を予測する予測モデルを作成する、作成ステップと、

前記予測モデルを、選択された前記説明変数を変数として含む第 2 のデータソースに適用して、疑似シングルソースデータを生成する、生成ステップと、

を有することを特徴とする情報処理方法。

【請求項 1 3】

請求項 1、2、10、11 および 12 のいずれか 1 項に記載の情報処理方法を情報処理装置に実行させるプログラム。

【請求項 1 4】

請求項 1、2、10、11 および 12 のいずれか 1 項に記載の情報処理方法を情報処理装置に実行させるプログラムを格納した、前記情報処理装置により読み取り可能な非一時的な記憶媒体。

【請求項 1 5】

請求項 1、2、10、11 および 12 のいずれか 1 項に記載の情報処理方法を実行する

30

20

40

50

、情報処理装置。

10

20

30

40

50