

(19)日本国特許庁(JP)

## (12)特許公報(B2)

(11)特許番号  
特許第7040104号  
(P7040104)

(45)発行日 令和4年3月23日(2022.3.23)

(24)登録日 令和4年3月14日(2022.3.14)

(51)国際特許分類		F I			
G 0 6 N	3/08 (2006.01)	G 0 6 N	3/08	1 4 0	
G 0 6 N	20/00 (2019.01)	G 0 6 N	20/00	1 3 0	

請求項の数 9 (全15頁)

(21)出願番号	特願2018-27256(P2018-27256)	(73)特許権者	000005223 富士通株式会社
(22)出願日	平成30年2月19日(2018.2.19)		神奈川県川崎市中原区上小田中4丁目1番1号
(65)公開番号	特開2019-144767(P2019-144767 A)	(74)代理人	110002147 特許業務法人酒井国際特許事務所
(43)公開日	令和1年8月29日(2019.8.29)	(72)発明者	梅田 裕平 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2年11月10日(2020.11.10)	審査官	吉倉 大智

最終頁に続く

(54)【発明の名称】 学習プログラム、学習方法および学習装置

## (57)【特許請求の範囲】

## 【請求項1】

コンピュータに、  
学習対象のデータそれぞれに付された1または複数のラベルそれぞれに対し、前記学習対象のデータの属性が複数の分布を含む混合分布に従う場合、前記混合分布における混合比率に基づき、スコアを設定し、  
前記学習対象のデータそれぞれに付されたラベルに設定されたスコアを用いて、ニューラルネットワークに対する学習を行わせる、  
処理を実行させる学習プログラム。

## 【請求項2】

コンピュータに、  
学習対象のデータそれぞれに付された1または複数のラベルそれぞれに対し、前記学習対象のデータから所定距離に位置する近接の学習対象のデータそれぞれを特定し、前記近接の学習対象のデータそれぞれに付されたラベルの割合に基づき、スコアを設定し、  
前記学習対象のデータそれぞれに付されたラベルに設定されたスコアを用いて、ニューラルネットワークに対する学習を行わせる、  
処理を実行させる学習プログラム。

## 【請求項3】

コンピュータに、  
学習対象のデータそれぞれに付された1または複数のラベルそれぞれに対し、前記学習対

象のデータから所定距離に位置する近接の学習対象のデータそれぞれを特定し、前記近接の学習対象のデータそれぞれに付されたラベルの割合と、前記学習対象のデータと前記近接の学習対象のデータそれぞれとの距離に応じた重みとを用いて、スコアを設定し、前記学習対象のデータそれぞれに付されたラベルに設定されたスコアを用いて、ニューラルネットワークに対する学習を行わせる、  
処理を実行させる学習プログラム。

【請求項 4】

コンピュータが、  
学習対象のデータそれぞれに付された 1 または複数のラベルそれぞれに対し、前記学習対象のデータの属性が複数の分布を含む混合分布に従う場合、前記混合分布における混合比率に基づき、スコアを設定し、  
前記学習対象のデータそれぞれに付されたラベルに設定されたスコアを用いて、ニューラルネットワークに対する学習を行わせる、  
処理を実行する学習方法。

10

【請求項 5】

コンピュータが、  
学習対象のデータそれぞれに付された 1 または複数のラベルそれぞれに対し、前記学習対象のデータから所定距離に位置する近接の学習対象のデータそれぞれを特定し、前記近接の学習対象のデータそれぞれに付されたラベルの割合に基づき、スコアを設定し、  
前記学習対象のデータそれぞれに付されたラベルに設定されたスコアを用いて、ニューラルネットワークに対する学習を行わせる、  
処理を実行する学習方法。

20

【請求項 6】

コンピュータが、  
学習対象のデータそれぞれに付された 1 または複数のラベルそれぞれに対し、前記学習対象のデータから所定距離に位置する近接の学習対象のデータそれぞれを特定し、前記近接の学習対象のデータそれぞれに付されたラベルの割合と、前記学習対象のデータと前記近接の学習対象のデータそれぞれとの距離に応じた重みとを用いて、スコアを設定し、  
前記学習対象のデータそれぞれに付されたラベルに設定されたスコアを用いて、ニューラルネットワークに対する学習を行わせる、  
処理を実行する学習方法。

30

【請求項 7】

学習対象のデータそれぞれに付された 1 または複数のラベルそれぞれに対し、前記学習対象のデータの属性が複数の分布を含む混合分布に従う場合、前記混合分布における混合比率に基づき、スコアを設定する設定部と、  
前記学習対象のデータそれぞれに付されたラベルに設定されたスコアを用いて、ニューラルネットワークに対する学習を行わせる学習部と、  
を有する学習装置。

【請求項 8】

学習対象のデータそれぞれに付された 1 または複数のラベルそれぞれに対し、前記学習対象のデータから所定距離に位置する近接の学習対象のデータそれぞれを特定し、前記近接の学習対象のデータそれぞれに付されたラベルの割合に基づき、スコアを設定する設定部と、  
前記学習対象のデータそれぞれに付されたラベルに設定されたスコアを用いて、ニューラルネットワークに対する学習を行わせる学習部と、  
を有する学習装置。

40

【請求項 9】

学習対象のデータそれぞれに付された 1 または複数のラベルそれぞれに対し、前記学習対象のデータから所定距離に位置する近接の学習対象のデータそれぞれを特定し、前記近接の学習対象のデータそれぞれに付されたラベルの割合と、前記学習対象のデータと前記近

50

接の学習対象のデータそれぞれとの距離に応じた重みとを用いて、スコアを設定する設定部と、

前記学習対象のデータそれぞれに付されたラベルに設定されたスコアを用いて、ニューラルネットワークに対する学習を行わせる学習部と、

を有する学習装置。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、学習プログラム、学習方法および学習装置に関する。

【背景技術】

【0002】

ラベル付きデータを用いた教師あり学習に関する技術が知られている。ラベル付きデータを用いた教師あり学習において用いられるラベルは、別の観点でデータの種別がはっきりした確実なラベルであることもあるが、作業者の主観により人手で付与されたラベルであることもある。一般的に、ラベル付きデータは、正解が既知である正解付きデータとして学習に用いられるから、正例と負例との境界付近のデータについてもいずれかのラベルを付与され、学習が行われる。

【0003】

図12は、一般的なラベル付与を説明する図である。図12の(a)に示すように、境界付近の曖昧なデータに対して、ラベルAとラベルBのいずれかを付与する必要がある場合、当該データの近傍のデータに付与されたラベルの多数決によって決定することが行われる。また、図12の(b)に示すように、境界付近の曖昧なデータに関しては、ラベルの確実性が低いので、学習データから除外することが行われる。

【先行技術文献】

【特許文献】

【0004】

【文献】特開2015-166962号公報

特開2017-016414号公報

【発明の概要】

【発明が解決しようとする課題】

【0005】

しかしながら、上記ラベルの付与方法では、学習結果の判定精度が劣化する場合がある。例えば、多数決を利用する手法では、ラベル付けが間違っていた場合、特に境界付近での誤差が大きくなる。また、ラベルが交じり合うことも多く、非線形性が高くなるので、判定器(分類器)の学習が難しい。除外する手法では、非線形性が低くなるので、学習はし易くなるが、境界付近が学習できないので、境界付近の判定精度が低下する。

【0006】

一つの側面では、学習結果の判定精度の劣化を抑制することができる学習プログラム、学習方法および学習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、学習プログラムは、コンピュータに、学習対象のデータそれぞれに付された1または複数のラベルそれぞれに対し、前記学習対象のデータの属性または前記学習対象のデータと他の学習対象のデータとの関係に基づき、スコアを設定する処理を実行させる。学習プログラムは、コンピュータに、前記学習対象のデータそれぞれに付されたラベルに設定されたスコアを用いて、ニューラルネットワークに対する学習を行わせる処理を実行させる。

【発明の効果】

【0008】

一実施形態によれば、学習結果の判定精度の劣化を抑制することができる。

10

20

30

40

50

## 【図面の簡単な説明】

【0009】

【図1】図1は、実施例1にかかる学習装置の全体例を説明する図である。

【図2】図2は、実施例1にかかる学習例を説明する図である。

【図3】図3は、実施例1にかかる学習装置の機能構成を示す機能ブロック図である。

【図4】図4は、学習データDBに記憶される情報の例を示す図である。

【図5】図5は、分布を用いたラベルの設定例を説明する図である。

【図6】図6は、近傍のデータの割合を用いたラベルの設定例を説明する図である。

【図7】図7は、データ間の距離を用いたラベルの設定例を説明する図である。

【図8】図8は、クラウドソーシングによるラベルの設定例を説明する図である。

10

【図9】図9は、処理の流れを示すフローチャートである。

【図10】図10は、効果を説明する図である。

【図11】図11は、ハードウェア構成例を説明する図である。

【図12】図12は、一般的なラベル付与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、本願の開示する学習プログラム、学習方法および学習装置の実施例を図面に基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。

## 【実施例1】

20

【0011】

## [全体構成]

図1は、実施例1にかかる学習装置の全体例を説明する図である。図1に示すように、実施例1にかかる学習装置10は、学習データのラベルに対してスコアを付与した後に、機械学習や深層学習（ディープラーニング（DL）・Deep Learning）などを用いた判別処理（学習処理）を実行して、学習データを事象ごとに正しく判別（分類）できるように、スコアを用いてニューラルネットワーク（NN：Neural Network）などを学習する。その後、学習結果を適用した学習モデルを用いることで、判別対象データの正確な事象（ラベル）の推定を実現する。なお、学習データには、画像、動画、文書、グラフなど様々なデータを採用することができる。

30

【0012】

例えば、学習装置10は、NNを用いたモデルに対する学習を行う際に、学習対象のデータそれぞれに付された1または複数のラベルそれぞれに対し、データの属性またはデータと他のデータとの関係に基づきスコアを設定する。そして、学習装置10は、学習対象のデータそれぞれに付されたラベルに設定されたスコアを用いて、NNに対する学習を行わせる。

【0013】

一般的に、NNの学習に各データに対して決定するラベルは、行列として保持する。しかし、従来から使用されていたSVM（Support Vector Machine）などのアルゴリズムが、ラベルを一つに決める必要があったこと、全学習データの認識スコアが正解ラベルに沿って1または0であることが最も望ましい状態であることから、複数のラベル成分に小数（分数）で値を設定することなく、1または0が設定されていた。

40

【0014】

つまり、1か0か曖昧なデータであっても1か0のいずれかを設定する必要があった。言い換えると、ラベルAかラベルBか曖昧なデータであっても、いずれかのラベルを設定する必要があるので、当該データに対して、ラベルとして「ラベル（ラベルA = 1.0、ラベルB = 0.0）」または「ラベル（ラベルA = 0.0、ラベルB = 1.0）」を付与する必要があった。

【0015】

そこで、実施例1では、ラベルが曖昧なデータについては、各ラベルである確率を各ラベ

50

ルに対応した要素に付したラベルベクトルを与え、これらのラベルベクトルをもとに深層学習を実行する。すなわち、実施例 1 では、付与されるラベルが曖昧なデータに対しては、確率的なラベルベクトルで与え、ラベルの値を小数として学習させる。

【 0 0 1 6 】

[ 学習例 ]

次に、ラベルが曖昧な学習データの学習について説明する。図 2 は、実施例 1 にかかる学習例を説明する図である。ここでは、図 2 の ( a ) と ( b ) は、一般的な学習例を示し、図 2 の ( c ) は、実施例 1 による学習例を示す。

【 0 0 1 7 】

図 2 の ( a ) に示すように、「ラベル A = 1 . 0、ラベル B = 0 . 0」が付与されたデータを NN に入力し、出力が、ラベル A である確率が 7 0 % でラベル B である確率が 3 0 % であったとする。この場合、誤差逆伝搬法によってラベル A と判別されるように NN の学習が実行されるが、学習データに設定されたラベルがある程度は正しかったことから、許容できる範囲内で正常に NN を学習することができる。

10

【 0 0 1 8 】

一方、図 2 の ( b ) に示すように、「ラベル A = 1 . 0、ラベル B = 0 . 0」が付与されたデータを NN に入力し、出力が、ラベル A である確率が 4 0 % でラベル B である確率が 6 0 % であったとする。この場合、学習データに設定されたラベルが間違っていた可能性が高いにも関わらず、誤差逆伝搬法によってラベル A と判別されるように NN の学習が実行され、間違った方向に NN が学習されるので、判別精度の劣化を引き起こす。

20

【 0 0 1 9 】

これに対して、図 2 の ( c ) に示すように、「ラベル A = 0 . 6、ラベル B = 0 . 4」が付与されたデータを NN に入力し、出力が、ラベル A である確率が 7 0 % でラベル B である確率が 3 0 % であったとする。この場合、誤差逆伝搬法によってラベル A と判別されるように NN の学習が実行されるが、学習データに設定されたラベルが正しかったことから、図 2 の ( a ) と比較してもより正常に NN を学習することができる。

【 0 0 2 0 】

このように、実施例 1 にかかる学習装置 1 0 は、ラベルが曖昧なデータに対しては、いずれかのラベルであると強制的に学習させるのではなく、曖昧さを残したまま、曖昧さを考慮した学習を実行することができる。したがって、学習装置 1 0 は学習結果の判定精度の劣化を抑制することができる。

30

【 0 0 2 1 】

[ 機能構成 ]

図 3 は、実施例 1 にかかる学習装置の機能構成を示す機能ブロック図である。図 3 に示すように、学習装置 1 0 は、通信部 1 1 と記憶部 1 2 と制御部 2 0 を有する。

【 0 0 2 2 】

通信部 1 1 は、他の装置との通信を制御する処理部であり、例えば通信インターフェースである。例えば、通信部 1 1 は、管理者の端末から、処理開始指示を受信する。また、通信部 1 1 は、管理者の端末等から、学習対象であるデータ（入力データ）を受信して入力データ DB 1 3 に格納する。

40

【 0 0 2 3 】

記憶部 1 2 は、プログラムやデータを記憶する記憶装置の一例であり、例えばメモリやハードディスクなどである。この記憶部 1 2 は、入力データ DB 1 3、学習データ DB 1 4、学習結果 DB 1 5 を記憶する。

【 0 0 2 4 】

入力データ DB 1 3 は、学習対象となる入力データを記憶するデータベースである。ここで記憶されるデータは、人手等によってラベルが設定されていてもよく、未設定であってもよい。なお、データは、管理者等によって格納することもでき、通信部 1 1 が受信して記憶することもできる。

【 0 0 2 5 】

50

学習データDB14は、学習対象の教師ありデータを記憶するデータベースである。具体的には、学習データDB14は、後述する制御部20によって、入力データDB13に記憶される入力データと当該入力データに設定されたラベルとを対応付けて記憶する。図4は、学習データDB14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、学習データDB14は、「データID、ラベル1、ラベル2、ラベル3」を対応付けて記憶する。図4の例では、データIDが「1」のデータには、「ラベル1、ラベル2、ラベル3」として「0.5、0、0.5」のラベルベクトルが設定されていることを示す。なお、ここで示したラベルベクトルの次元数や数値は、一例であり、任意に設定変更することができる。

#### 【0026】

学習結果DB15は、学習結果を記憶するデータベースである。例えば、学習結果DB15は、制御部20による学習データの判別結果（分類結果）、機械学習やディープラーニングによって学習された各種パラメータを記憶する。

#### 【0027】

制御部20は、学習装置10全体の処理を司る処理部であり、例えばプロセッサなどである。この制御部20は、設定部21と学習部22を有する。なお、設定部21と学習部22は、プロセッサなどが有する電子回路やプロセッサなどが実行するプロセスの一例である。

#### 【0028】

設定部21は、学習対象のデータそれぞれに付された1または複数のラベルそれぞれに対し、データの属性またはデータと他のデータとの関係に基づきスコアを設定する処理部である。具体的には、設定部21は、入力データDB13から各入力データを読み出し、各入力データに基づきスコアを算出する。そして、設定部21は、各入力データに対して、スコアを設定したラベルベクトルをラベルとして設定した学習データを生成する。その後、設定部21は、生成した学習データを学習データDB14に格納する。なお、入力データに人手等によるラベルが既に付与されている場合は、ラベルの補正が行われる。また、後述する処理によって、曖昧なデータについてだけラベルの再設定を行うこともでき、全データについてラベルの再設定を行うこともできる。

#### 【0029】

つまり、設定部21は、NNの学習において、ラベル付けされたデータの「確信度や信頼度」が「全て正しい」という前提をあてはめることによる弊害を、小数ラベル（ラベルベクトル）により解決する。ここで、設定部21が実行するラベルの設定方法の具体例について説明する。なお、ラベルが2つ（二次元）の場合を用いて説明するが、これに限定されるものではなく、三次元以上であっても同様に処理することができる。なお、一例としては、設定部21は、複数の管理者等のユーザによってラベルが異なるデータを曖昧なデータと判定することができる。

#### 【0030】

（手法1：分布）

まず、曖昧なデータの属性が複数の分布を含む混合分布に従う場合、混合分布における混合比率に基づきスコアを設定する例を説明する。つまり、各ラベルの発生がある分布に沿っていると仮定し、各ラベルの混合分布に基づいて決定する手法を説明する。この例では、各データ間の距離が定まっており、データ数は十分に存在し、曖昧なラベルも含めて全データにラベルが付与されているものとする。

#### 【0031】

図5は、分布を用いたラベルの設定例を説明する図である。この例は、同年代の身長と体重の数値から男女を識別する例である。身長と体重はセンサーで測定され、ラベル付けは目視で行う場合もしくは分布に沿って自動的に行う場合を考える。図5に示すように、正規化を行った身長と体重の分布は、正規分布に従うことが予想され、男性の方が身長および体重の平均が大きい。

#### 【0032】

10

20

30

40

50

図5の例では、女性の正規分布のみに沿うデータを丸印、男性の正規分布のみに沿うデータを点線の丸印で表す。例えば、設定部21は、正規分布が重複しない領域かつ女性の正規分布に属する領域のデータ( $ID = 1$ )に対しては、ラベルベクトル「ラベル1(女性) = 1.0、ラベル2(男性) = 0.0」を設定する。また、設定部21は、正規分布が重複しない領域かつ男性の正規分布に属する領域のデータ( $ID = 20$ )に対しては、ラベルベクトル「ラベル1(女性) = 0.0、ラベル2(男性) = 1.0」を設定する。

#### 【0033】

これらに対して、設定部21は、分布が重複する領域Pに属するデータ( $ID = D$ )、すなわち曖昧なデータDに対しては、混合分布の割合等に基づくスコアをラベルとして設定する。例えば、設定部21は、女性分布上の値P2と男性分布上の値P1とを特定し、P0からP1までの距離( $P1 - P0$ )と、P0からP2までの距離( $P2 - P0$ )との割合を算出する。そして、設定部21は、「距離( $P2 - P0$ ) : 距離( $P1 - P0$ )」 = 「6 : 4」と算出した場合、データDに対しては、ラベルベクトル「ラベル1(女性) = 0.6、ラベル2(男性) = 0.4」を設定する。

10

#### 【0034】

なお、設定部21は、両方の分布に属する各データ、言い換えると両方の分布に沿うデータを曖昧なデータと判定し、上記処理によってスコアを算出する。なお、割合の算出時は、合計が1になるように正規化することもできる。また、距離に限らず、値そのもの(図5では身長)の割合や比率などを用いることもできる。また、いずれかの分布に沿っているデータについては管理者等による人手でラベルを設定し、曖昧なデータのみに対して上記手法1によるラベル設定を実行することもできる。

20

#### 【0035】

(手法2:近傍データの割合)

次に、曖昧なデータの近傍のデータに付与されるラベルの割合に基づいて、曖昧なデータにラベルを設定する例を説明する。この例でも手法1と同様、各データ間の距離が定まっており、データ数は十分に存在し、曖昧なラベルも含めて全データにラベルが付与されているものとする。なお、データが三次元以上の場合には、全データ間の距離を計算し、MDS(Multi-Dimensional Scaling)などで二次元に次元圧縮する。

#### 【0036】

図6は、近傍のデータの割合を用いたラベルの設定例を説明する図である。この例は、機器の動作時の各部位の振動から、機器が正常か異常か判断し、各部位の振動データである各データに対して正常か異常などのラベルを設定する例である。機器の異常は経年劣化的に起こることから、正常と異常の境目の判断は不確実性が高い。また、境界付近で判断が曖昧であることが多く、正常と異常のデータはそれぞれ分布に従っているわけではない。

30

#### 【0037】

図6の例では、過去の事例や実際に発生した障害事例などから、正常値と判定されたデータを丸印、異常値と判定されたデータを点線の丸印で表す。例えば、設定部21は、正常値と判定されたデータ( $ID = 1$ )に対しては、ラベルベクトル「ラベル1(正常) = 1.0、ラベル2(異常) = 0.0」を設定する。また、設定部21は、異常値と判定されたデータ( $ID = 20$ )に対しては、ラベルベクトル「ラベル1(正常) = 0.0、ラベル2(異常) = 1.0」を設定する。

40

#### 【0038】

これらに対して、設定部21は、過去の事例等より、正常値か異常値かの判断ができない曖昧なデータ( $ID = D$ )に対しては、圧縮空間上である閾値の距離内の近傍に存在する他のデータのラベルの割合に基づいてラベル設定を行う。なお、図6の丸印内の数字はデータIDを示す。

#### 【0039】

図6に示すように、設定部21は、MDSなどで得られるデータ間の距離を用いて、曖昧なデータDから任意の所定範囲Q内に存在するデータを特定する。そして、設定部21は、所定範囲Q内のデータのうち、データ1、3、5、10の4つのデータのラベルが「正

50

常」であり、データ 2、4、6、7、8、9 の 6 つのデータのラベルが「異常」であることを特定する。すなわち、設定部 21 は、所定範囲 Q 内の近傍データの 4 割が「正常」で 6 割が「異常」と特定済みであると判定する。この結果、設定部 21 は、データ D に対しては、ラベルベクトル「ラベル 1 (正常) = 0.4、ラベル 2 (異常) = 0.6」を設定する。

#### 【0040】

なお、設定部 21 は、管理者等のユーザによって正常か異常かの区別ができないと判定されたデータ、過去事例に基づき正常にも異常にも属しないと判定されたデータなどを曖昧なデータと判定することができる。なお、割合の算出時は、合計が 1 になるように正規化することもできる。また、正常か異常かを正確に判定されたデータについては管理者等による人手でラベルを設定し、曖昧なデータのみに対して上記手法 2 によるラベル設定を実行することもできる。

10

#### 【0041】

(手法 3 : データ間の距離)

次に、曖昧なデータの近傍のデータ間の距離に基づいて、曖昧なデータにラベルを設定する例を説明する。この例の条件は、手法 2 と同様とする。図 7 は、データ間の距離を用いたラベルの設定例を説明する図である。

#### 【0042】

図 7 に示すように、設定部 21 は、MDS など得られるデータ間の距離を用いて、曖昧なデータ D から任意の所定範囲 Q 内に存在するデータを特定する。そして、設定部 21 は、所定範囲 Q 内のデータのうち、「正常」と識別された (正常のラベルのみが付与された) データ 1、3、5、10 の 4 つのデータを特定する。続いて、設定部 21 は、予め算出済みのデータ間の距離を用いて、データ D とデータ 1 との距離  $w_1$ 、データ D とデータ 3 との距離  $w_3$ 、データ D とデータ 5 との距離  $w_5$ 、データ D とデータ 10 との距離  $w_{10}$  を算出する。その後、設定部 21 は、距離による重み ( $w$  の総和) として、「 $(1/w_1) + (1/w_3) + (1/w_5) + (1/w_{10})$ 」を算出する。なお、ここでは重みの算出に距離の逆数を用いたが、距離が近いほど大きくなる指標であれば逆数以外の指標を用いてもよい。

20

#### 【0043】

同様に、設定部 21 は、所定範囲 Q 内のデータのうち、「異常」と識別された (異常のラベルのみが付与された) データ 2、4、6、7、8、9 の 6 つのデータを特定する。続いて、設定部 21 は、予め算出済みのデータ間の距離を用いて、データ D とデータ 2 との距離  $W_2$ 、データ D とデータ 4 との距離  $W_4$ 、データ D とデータ 6 との距離  $W_6$ 、データ D とデータ 7 との距離  $W_7$ 、データ D とデータ 8 との距離  $W_8$ 、データ D とデータ 9 との距離  $W_9$  を算出する。その後、設定部 21 は、距離による重み ( $W$  の総和) として、「 $(1/W_2) + (1/W_4) + (1/W_6) + (1/W_7) + (1/W_8) + (1/W_9)$ 」を算出する。

30

#### 【0044】

この結果、設定部 21 は、データ D に対しては、ラベルベクトル「ラベル 1 (正常)、ラベル 2 (異常)」として「ラベル 1 (正常) =  $w$  の総和、ラベル 2 (異常) =  $W$  の総和」を設定する。なお、距離の重みを考慮した算出手法は、一例であり、距離が近いほど重要視されるような手法であれば、どのような手法も採用することができる。また、距離による重みは、合計が 1 になるように正規化して算出することができる。また、手法 2 と手法 3 では、全データについて、上記で計算した確率 (値) はなめらかな関数にならないため、各ラベルに対し応答曲面を作成し、各ラベルの応答曲面による値をベクトルのセル値に対応させることもできる。

40

#### 【0045】

(手法 4 : 近傍データの割合)

次に、ラベル決定の参考となる情報が複数存在する場合に、参考情報で指示されるラベルの割合に基づいて、ラベルを設定する例を説明する。例えば、クラウドソーシング等によ

50

って複数の担当者にラベル付作業を依頼することが考えられる。この場合、それぞれのラベル付結果から各データのラベルを決定するが、曖昧なデータに関しては、各担当者によって付与されるラベルが異なる場合がある。

【 0 0 4 6 】

一般的には、多数決や担当者の信頼度によって決定されるが、必ずしも正しいラベルが付与されるとは限らない。そこで、設定部 2 1 は、ラベル付結果の割合に基づいてラベルベクトルを生成して設定する。

【 0 0 4 7 】

図 8 は、クラウドソーシングによるラベルの設定例を説明する図である。図 8 に示すように、データ D に対して、担当者 a がラベル 1 を付与し、担当者 b がラベル 1 を付与し、担当者 c がラベル 1 を付与し、担当者 d がラベル 2 を付与し、担当者 e がラベル 1 を付与したとする。この場合、設定部 2 1 は、各ラベルの設定合計数を計算し、ラベル 1 が「4」、ラベル 2 が「1」と算出する。そして、設定部 2 1 は、各ラベルの全体に対する割合「ラベル 1、ラベル 2」として「 $4 / 5 = 0.8$ 、 $1 / 5 = 0.2$ 」を算出する。この結果、設定部 2 1 は、データ D に対しては、ラベルベクトル「ラベル 1 = 0.8、ラベル 2 = 0.2」を設定する。

【 0 0 4 8 】

なお、担当者の信頼度等によって、重み付けすることもできる。例えば、予め指定された担当者 a の信頼度が閾値以上である場合に、担当者 a については、設定数が 1 つであっても 2 倍にして 2 として上記算出手法を実行することもできる。また、各参考情報で、指示するラベルが異なる場合、参考情報の重要度に応じて重みをつけ、各ラベルを指示する情報の重み付和を全体の重み付和で割った「各ラベルの重み付比率」を各ラベルに対する値とすることもできる。

【 0 0 4 9 】

図 3 に戻り、学習部 2 2 は、学習データ DB 1 4 に記憶される学習データを用いて NN の学習を実行し、学習結果を学習結果 DB 1 5 に格納する処理部である。図 4 の例では、学習部 2 2 は、ID = 1 のデータについてはラベルベクトル「ラベル 1 = 0.5、ラベル 2 = 0、ラベル 3 = 0.5」を入力として学習を実行する。

【 0 0 5 0 】

[ 処理の流れ ]

次に、上述したラベルベクトルの設定処理について説明する。図 9 は、処理の流れを示すフローチャートである。

【 0 0 5 1 】

図 9 に示すように、入力データが受信されて入力データ DB 1 3 に格納されると ( S 1 0 1 : Y e s )、設定部 2 1 は、入力データ DB 1 3 から 1 つの入力データを読み込む ( S 1 0 2 )。

【 0 0 5 2 】

続いて、設定部 2 1 は、読み込んだ入力データが曖昧なデータに該当するか否かを判定し ( S 1 0 3 )、曖昧なデータに該当する場合 ( S 1 0 3 : Y e s )、入力データの属性や他データとの関係からスコアを算出する ( S 1 0 4 )。そして、設定部 2 1 は、スコアに基づいたラベルベクトルを入力データに設定 ( 付与 ) した学習データを生成して ( S 1 0 5 )、学習データ DB 1 4 に格納する ( S 1 0 6 )。

【 0 0 5 3 】

一方、設定部 2 1 は、読み込んだ入力データが曖昧なデータに該当しない場合 ( S 1 0 3 : N o )、入力データに対して既知のラベルを示すラベルベクトルを設定した学習データを生成して ( S 1 0 7 )、学習データ DB 1 4 に格納する ( S 1 0 6 )。なお、曖昧ではない入力データにすでに付与されているラベルをそのまま用いることができる。

【 0 0 5 4 】

その後、全入力データに対してラベル ( ラベルベクトル ) が設定済みではなく、未設定の入力データが存在する場合 ( S 1 0 8 : N o )、S 1 0 2 以降が実行される。

10

20

30

40

50

## 【0055】

一方、全入力データに対してラベル（ラベルベクトル）が設定済みである場合（S108：Yes）、学習部22は、学習データDB14から各学習データを読み込み（S109）、各学習データのラベルベクトルをもとに学習を実行する（S110）。

## 【0056】

## [効果]

上述したように、学習装置10は、付与されたラベルが曖昧な場合に、確率的なラベルベクトルを与えることで深層学習を行い精度の高い学習を行うことができる。また、学習装置10は、ラベルを集約させることによる、判別速度の劣化や学習結果の判別精度の劣化を抑制することができる。

10

## 【0057】

ここで、実施例1による手法と従来手法との比較を行った実験結果を説明する。まず、実験の条件を説明する。ここでは、10次元のベクトルデータを用い、第1成分が0.5以上か否かで正例か負例か分類する例で説明する。なお、曖昧なデータの条件として、第1成分が0.35から0.55の間のデータでは3割の確率でランダムにラベルを入れ替えることとする。

## 【0058】

比較する手法は、そのままのラベルで学習する手法「一般法1」と、担当者の主観によってラベルを付け替えた「一般法2」、不確定な区間である（0.35から0.6）区間のデータを学習データから除去した「不確定除去」と、上記手法1から4のいずれかを用いた「実施例1」とする。

20

## 【0059】

図10は、効果を説明する図である。図10では、各手法で学習データを生成した上で学習を実行し、その後、学習結果を反映した学習モデルを用いて、判別対象データの判別を実行した結果を示す。図10に示すように、全体的な精度は、各手法とも精度の高い判別（分類）ができたが、不確定な範囲（0.35から0.6の区間）については、各手法とも精度が低下した。しかし、実施例1では、精度が低下したものの、依然として80%以上の精度を維持しており、高精度に判別できたことがわかる。したがって、実施例1は、他の手法と比較しても、劣化や学習結果の判別精度の劣化を抑制することができる。

## 【実施例2】

30

## 【0060】

さて、これまで本発明の実施例について説明したが、本発明は上述した実施例以外にも、種々の異なる形態にて実施されてよいものである。

## 【0061】

## [システム]

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することができる。また、実施例で説明した具体例、分布、数値などは、あくまで一例であり、任意に変更することができる。

## 【0062】

40

また、図示した各装置の各構成要素は機能概念的なものであり、必ずしも物理的に図示の如く構成されていることを要しない。すなわち、各装置の分散や統合の具体的形態は図示のものに限られない。つまり、その全部または一部を、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・統合して構成することができる。さらに、各装置にて行なわれる各処理機能は、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。

## 【0063】

## [ハードウェア]

図11は、ハードウェア構成例を説明する図である。図11に示すように、学習装置10

50

は、通信装置 10 a、HDD (Hard Disk Drive) 10 b、メモリ 10 c、プロセッサ 10 d を有する。また、図 11 に示した各部は、バス等で相互に接続される。

【0064】

通信装置 10 a は、ネットワークインタフェースカードなどであり、他のサーバとの通信を行う。HDD 10 b は、図 3 に示した機能を動作させるプログラムや DB を記憶する。

【0065】

プロセッサ 10 d は、図 3 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 10 b 等から読み出してメモリ 10 c に展開することで、図 3 等で説明した各機能を実行するプロセスを動作させる。すなわち、このプロセスは、学習装置 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、設定部 21、学習部 22 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、設定部 21、学習部 22 等と同様の処理を実行するプロセスを実行する。

10

【0066】

このように学習装置 10 は、プログラムを読み出して実行することで学習方法を実行する情報処理装置として動作する。また、学習装置 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、この他の実施例でいうプログラムは、学習装置 10 によって実行されることに限定されるものではない。例えば、他のコンピュータまたはサーバがプログラムを実行する場合や、これらが協働してプログラムを実行するような場合にも、本発明を同様に適用することができる。

20

【符号の説明】

【0067】

10 学習装置

11 通信部

12 記憶部

13 入力データ DB

14 学習データ DB

15 学習結果 DB

20 制御部

21 設定部

22 学習部

30

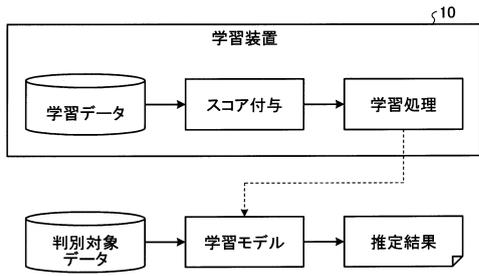
40

50

【 図 面 】

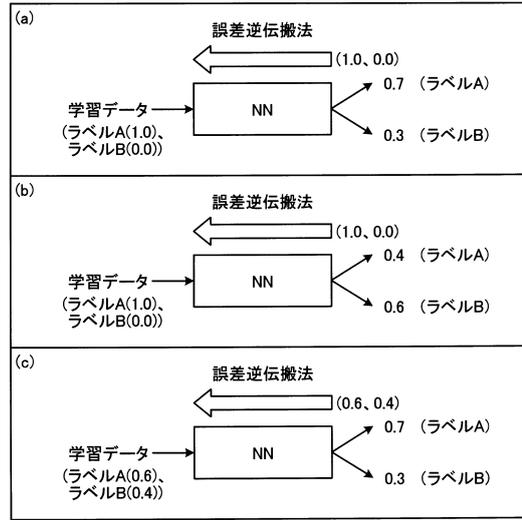
【 図 1 】

実施例1にかかる学習装置の全体例を説明する図



【 図 2 】

実施例1にかかる学習例を説明する図

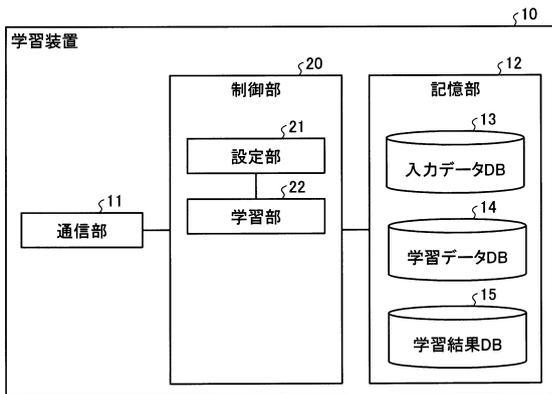


10

20

【 図 3 】

実施例1にかかる学習装置の機能構成を示す機能ブロック図



【 図 4 】

学習データDBに記憶される情報の例を示す図

データID	ラベル1	ラベル2	ラベル3
1	0.5	0	0.5
2	0	1.0	0
...	...	...	...

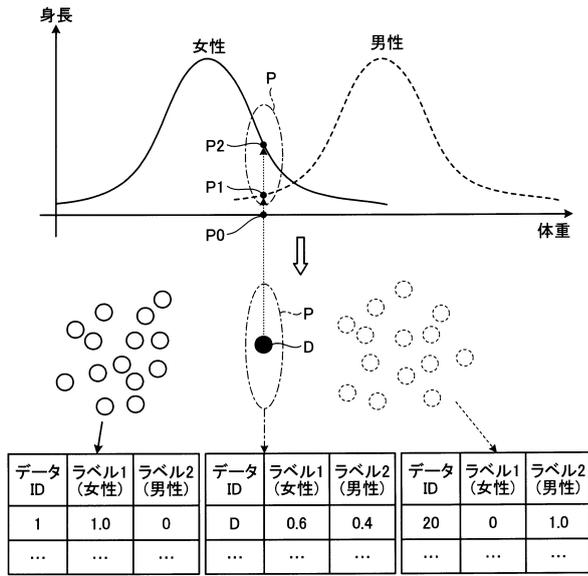
30

40

50

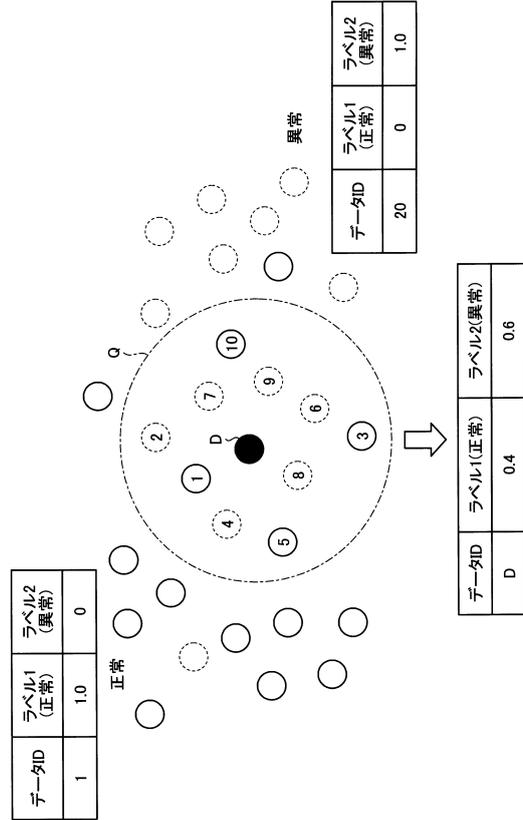
【 図 5 】

分布を用いたラベルの設定例を説明する図



【 図 6 】

近傍のデータの割合を用いたラベルの設定例を説明する図

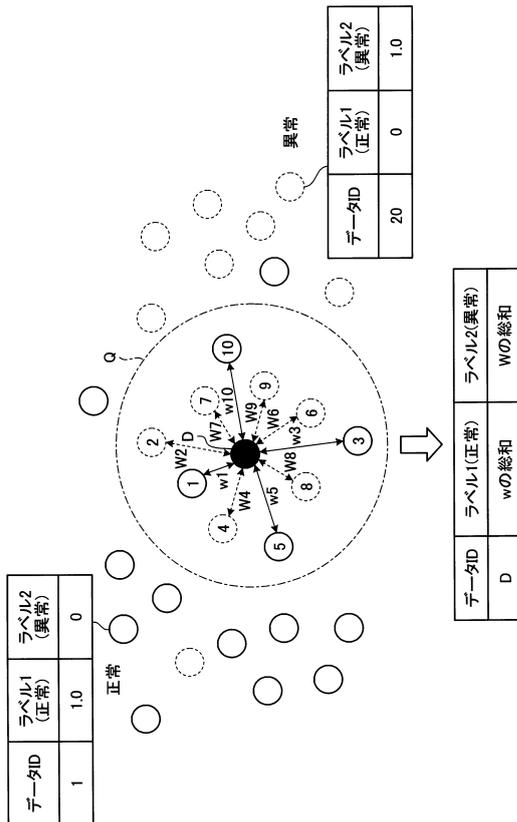


10

20

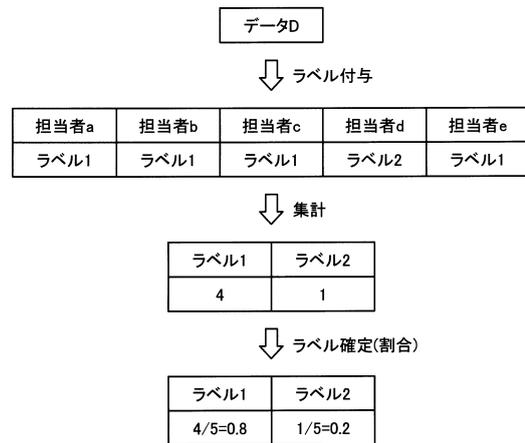
【 図 7 】

データ間の距離を用いたラベルの設定例を説明する図



【 図 8 】

クラウドソーシングによるラベルの設定例を説明する図

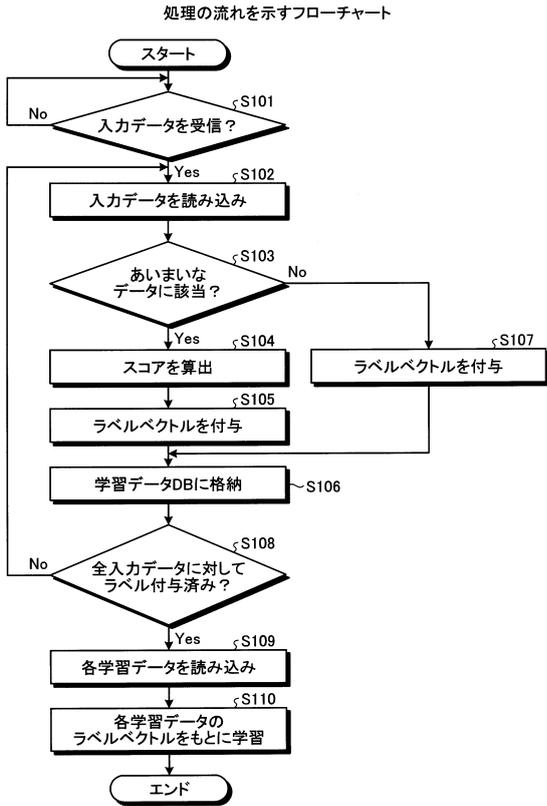


30

40

50

【 図 9 】



【 図 1 0 】

効果を説明する図

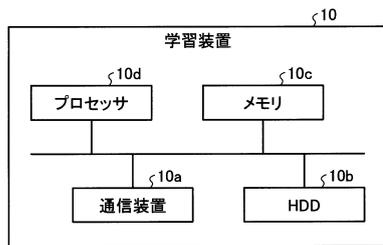
	一般法1	一般法2	不確定除去	実施例1
精度(全体)	86%	91%	92%	96%
精度(不確定範囲)	56%	74%	70%	84%

10

20

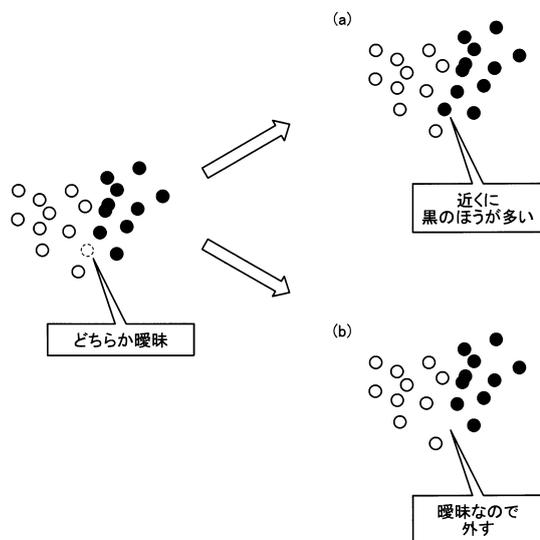
【 図 1 1 】

ハードウェア構成例を説明する図



【 図 1 2 】

一般的なラベル付与を説明する図



30

40

50

---

フロントページの続き

- (56)参考文献 特開2009-282686(JP,A)  
国際公開第2017/073373(WO,A1)  
国際公開第2017/183242(WO,A1)  
特表2016-505974(JP,A)  
特開2009-110064(JP,A)  
得居誠也, 最適化から見たディープラーニングの考え方, オペレーションズ・リサーチ, 日本, 公益社団法人日本オペレーションズ・リサーチ学会, 2015年04月01日, 第60巻 第4号, PP.191-197, ISSN: 0030-3674
- (58)調査した分野 (Int.Cl., DB名)  
G06N 3/00-99/00