

(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号
特許第7287505号
(P7287505)

(45)発行日 令和5年6月6日(2023.6.6)

(24)登録日 令和5年5月29日(2023.5.29)

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

請求項の数 7 (全27頁)

(21)出願番号	特願2021-570583(P2021-570583)	(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
(86)(22)出願日	令和2年1月16日(2020.1.16)	(74)代理人	110002147 弁理士法人酒井国際特許事務所
(86)国際出願番号	PCT/JP2020/001382	(72)発明者	溝淵 裕司 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
(87)国際公開番号	WO2021/144938	審査官	千葉 久博
(87)国際公開日	令和3年7月22日(2021.7.22)		
審査請求日	令和4年4月18日(2022.4.18)		

最終頁に続く

(54)【発明の名称】 学習支援プログラム、学習支援方法及び学習支援装置

(57)【特許請求の範囲】

【請求項1】

学習データに含まれるサンプルがクラスタリングされたクラスタごとに、前記クラスタを代表するサンプルを第1の距離計量モデルへ入力することにより第1の埋め込みベクトルを算出し、

前記学習データに含まれるサンプルのペアに設定されたラベルに基づいて前記第1の距離計量モデルから第2の距離計量モデルを学習し、

前記クラスタごとに、前記クラスタを代表するサンプルを前記第2の距離計量モデルへ入力することにより第2の埋め込みベクトルを算出し、

各クラスタの前記第1の埋め込みベクトルおよび各クラスタの前記第2の埋め込みベクトルに基づいて、前記第2の距離計量モデルの学習時のエポック数よりも多いエポック数で学習が行われる場合に統合され得るクラスタのペアを検出し、

前記クラスタのペアのうち類似のラベルが設定されていないクラスタのペアを出力する、処理をコンピュータに実行させる学習支援プログラム。

【請求項2】

前記検出する処理は、前記第2の距離計量モデルが学習される前後におけるクラスタ間の進行角度の大きさに基づいて前記クラスタのペアを検出する請求項1に記載の学習支援プログラム。

【請求項3】

前記検出する処理は、前記第2の距離計量モデルが学習される前後におけるクラスタ間

10

20

の距離の変化量に基づいて前記クラスタのペアを検出する請求項 1 に記載の学習支援プログラム。

【請求項 4】

前記第 2 の距離計量モデルが学習された後におけるクラスタ間の距離を、前記第 2 の距離計量モデルが学習される前におけるクラスタ間の距離で除算することにより、前記クラスタ間の距離の変化量を算出する処理を前記コンピュータにさらに実行させる請求項 3 に記載の学習支援プログラム。

【請求項 5】

前記出力する処理は、前記類似のラベルが設定されていないクラスタのペアに非類似のラベルを設定するレコメンドを出力する請求項 1 に記載の学習支援プログラム。

10

【請求項 6】

学習データに含まれるサンプルがクラスタリングされたクラスタごとに、前記クラスタを代表するサンプルを第 1 の距離計量モデルへ入力することにより第 1 の埋め込みベクトルを算出し、

前記学習データに含まれるサンプルのペアに設定されたラベルに基づいて前記第 1 の距離計量モデルから第 2 の距離計量モデルを学習し、

前記クラスタごとに、前記クラスタを代表するサンプルを前記第 2 の距離計量モデルへ入力することにより第 2 の埋め込みベクトルを算出し、

各クラスタの前記第 1 の埋め込みベクトルおよび各クラスタの前記第 2 の埋め込みベクトルに基づいて、前記第 2 の距離計量モデルの学習時のエポック数よりも多いエポック数で学習が行われる場合に統合され得るクラスタのペアを検出し、

20

前記クラスタのペアのうち類似のラベルが設定されていないクラスタのペアを出力する、処理をコンピュータが実行する学習支援方法。

【請求項 7】

学習データに含まれるサンプルがクラスタリングされたクラスタごとに、前記クラスタを代表するサンプルを第 1 の距離計量モデルへ入力することにより第 1 の埋め込みベクトルを算出する第 1 算出部と、

前記学習データに含まれるサンプルのペアに設定されたラベルに基づいて前記第 1 の距離計量モデルから第 2 の距離計量モデルを学習する学習部と、

前記クラスタごとに、前記クラスタを代表するサンプルを前記第 2 の距離計量モデルへ入力することにより第 2 の埋め込みベクトルを算出する第 2 算出部と、

30

各クラスタの前記第 1 の埋め込みベクトルおよび各クラスタの前記第 2 の埋め込みベクトルに基づいて、前記第 2 の距離計量モデルの学習時のエポック数よりも多いエポック数で学習が行われる場合に統合され得るクラスタのペアを検出する検出部と、

前記クラスタのペアのうち類似のラベルが設定されていないクラスタのペアを出力する出力部と、

を有することを特徴とする学習支援装置。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、学習支援プログラム、学習支援方法及び学習支援装置に関する。

40

【背景技術】

【0002】

距離計量学習の例として、マハラノビス距離学習をはじめ、ディープラーニング技術を利用したディープメトリックラーニングなどが提案されている。

【先行技術文献】

【特許文献】

【0003】

【文献】特開 2010 - 198518 号公報

【非特許文献】

50

【 0 0 0 4 】

【文献】Weinberger, Kilian Q., John Blitzer, and Lawrence K.Saul. "Distance metric learning for large margin nearest neighbor classification." Advances in neural information processing systems. 2006.

Gao, Xingyu, et al. "SOML: Sparse Online Metric Learning with Application to Image Retrieval." AAAI. 2014.

E.P.Xing, A.Y.Ng, M.I.Jordan and S.Russell: "Distance metric learning, with application to clustering with side-information", Neural Information Processing Systems(NIPS)(2003).

【発明の概要】

10

【発明が解決しようとする課題】

【 0 0 0 5 】

しかしながら、上記の距離計量学習が行われたモデルの再学習時には、学習サンプルのデータ間に設定されるラベルが再学習後のモデルに悪影響を与えることがある。

【 0 0 0 6 】

1つの側面では、本発明は、再学習後の距離計量モデルが悪影響を受けるのを抑制できる学習支援プログラム、学習支援方法及び学習支援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】

【 0 0 0 7 】

一態様の学習支援プログラムは、学習データに含まれるサンプルがクラスタリングされたクラスタごとに、前記クラスタを代表するサンプルを第1の距離計量モデルへ入力することにより第1の埋め込みベクトルを算出し、前記学習データに含まれるサンプルのペアに設定されたラベルに基づいて前記第1の距離計量モデルから第2の距離計量モデルを学習し、前記クラスタごとに、前記クラスタを代表するサンプルを前記第2の距離計量モデルへ入力することにより第2の埋め込みベクトルを算出し、各クラスタの前記第1の埋め込みベクトルおよび各クラスタの前記第2の埋め込みベクトルに基づいて、前記第2の距離計量モデルの学習時のエポック数よりも多いエポック数で学習が行われる場合に統合され得るクラスタのペアを検出し、前記クラスタのペアのうち類似のラベルが設定されていないクラスタのペアを出力する、処理をコンピュータに実行させる。

20

【発明の効果】

30

【 0 0 0 8 】

再学習後の距離計量モデルが悪影響を受けるのを抑制できる。

【図面の簡単な説明】

【 0 0 0 9 】

【図1】図1は、実施例1に係るシステムの構成例を示す図である。

【図2】図2は、多クラス分類モデルの一例を示す図である。

【図3】図3は、Siamese Networkの構造の一例を示す図である。

【図4】図4は、距離空間の一例を示す図である。

【図5】図5は、距離空間の一例を示す図である。

【図6】図6は、文書データの一例を示す図である。

40

【図7】図7は、単語列の抽出結果の一例を示す図である。

【図8】図8は、ラベル設定画面の一例を示す図である。

【図9】図9は、埋め込みベクトルの一例を示す図である。

【図10】図10は、埋め込みベクトルの一例を示す図である。

【図11】図11は、問合せ画面の一例を示す図である。

【図12】図12は、ラベル設定画面の一例を示す図である。

【図13】図13は、実施例1に係る学習支援処理の手順を示すフローチャートである。

【図14】図14は、コンピュータのハードウェア構成例を示す図である。

【発明を実施するための形態】

【 0 0 1 0 】

50

以下に添付図面を参照して本願に係る学習支援プログラム、学習支援方法及び学習支援装置について説明する。なお、この実施例は開示の技術を限定するものではない。そして、各実施例は、処理内容を矛盾させない範囲で適宜組み合わせることが可能である。

【実施例 1】

【0011】

[1 . システム構成]

図 1 は、実施例 1 に係るシステムの構成例を示す図である。図 1 に示すシステム 1 は、あくまで 1 つの側面として、任意の機械学習のタスク、例えばクラス分類に対応する機能を提供するものである。

【0012】

図 1 に示すように、システム 1 には、サーバ装置 10 と、クライアント端末 30 とが含まれ得る。これらサーバ装置 10 及びクライアント端末 30 は、ネットワーク NW を介して通信可能に接続される。例えば、ネットワーク NW は、有線または無線を問わず、インターネットや LAN (Local Area Network) などの任意の種類通信網であってかまわない。なお、図 1 には、1 つのサーバ装置 10 につき 1 つのクライアント端末 30 が接続される例を挙げたが、これはあくまで模式図に過ぎず、1 つのサーバ装置 10 につき複数のクライアント端末 30 が接続されることを妨げない。

【0013】

サーバ装置 10 は、上記のクラス分類を実行するコンピュータの一例である。サーバ装置 10 は、学習支援装置に対応し得る。

【0014】

一実施形態として、サーバ装置 10 は、上記のクラス分類に対応する機能を実現する分類プログラムを任意のコンピュータにインストールさせることによって実装できる。例えば、サーバ装置 10 は、上記のクラス分類に関する機能をオンプレミスに提供するサーバとして実装することができる。この他、サーバ装置 10 は、SaaS (Software as a Service) 型のアプリケーションとして実装することで、上記のクラス分類に対応する機能をクラウドサービスとして提供することとしてもかまわない。

【0015】

クライアント端末 30 は、上記のクラス分類に対応する機能の提供を受けるコンピュータの一例に対応する。例えば、クライアント端末 30 には、パーソナルコンピュータなどのデスクトップ型のコンピュータなどが対応する。これはあくまで一例に過ぎず、クライアント端末 30 は、ラップトップ型のコンピュータや携帯端末装置、ウェアラブル端末などの任意のコンピュータであってかまわない。

【0016】

[2 . タスクの実装例]

上記のクラス分類のタスクが実装され得る機能の例として、2 つの文書間の類似または非類似を判別する文書間の類似判別機能が挙げられる。あくまで一例として、上記の文書間の類似判別機能は、IT (Information Technology) サービス等の運用管理に関する過去の事例集からトラブル事象及びトラブル原因の間で類似する事例サンプルが関連付けられた障害切り分けグラフの構築を可能とする。このように構築される障害切り分けグラフは、あくまで一例として、トラブルへの対処時にトラブル事象に対応するトラブル原因のレコメンドを出力する機能等を実現し得る。

【0017】

[3 . 背景技術]

このような文書間の類似判別機能等で例示されるタスクにおけるクラス分類の判別精度を高める側面から、クラス分類には、特徴量の重要度に基づく特徴量選択や特徴量空間を学習する方法といった技術が組み込まれることがある。

【0018】

[3 . 1 マハラノビス距離学習]

特徴量空間を学習する方法の一例として、マハラノビス距離学習に代表される距離計量

10

20

30

40

50

学習が知られている。いわゆる距離計量学習では、入力空間におけるサンプル間の類似度を特徴量空間における距離に対応させる変換が学習される。すなわち、距離計量学習では、同じクラスに属するサンプル同士の距離が近く、異なるクラスに属するサンプル同士の距離が遠くなるように元の空間が歪められる。なお、「特徴量空間」は、距離空間、あるいは埋め込み空間と呼ばれる場合もある。

【 0 0 1 9 】

例えば、マハラノビス距離学習では、下記の式 (1) に示すマハラノビス距離 $d_M(x, x')$ を定義し、 M の成分を設計変数として学習が行われる。このような M の最適化問題は、サンプル間のユークリッド距離をサンプル間の類似度に対応させる変換 L を学習することと等価である。このことから、下記の式 (2) のロス関数を最小化する問題を解くことで、変換 L を学習することができる。

10

【 0 0 2 0 】

【数 1】

$$d_M(x, x') = \sqrt{(x - x')^T M (x - x')} \quad \dots (1)$$

【数 2】

$$\varepsilon(L) = (1 - \mu)\varepsilon_{pull}(L) + \mu\varepsilon_{push}(L). \quad \dots (2)$$

20

【 0 0 2 1 】

上記の式 (2) における「 L 」は、入力空間から線形変換する行列であり、 $M = L^T L$ の関係を有する。上記の式 (2) に含まれる 2 つのロス (ペナルティ) 関数のうち、 $\varepsilon_{pull}(L)$ は、下記の式 (3) で表されると共に、 $\varepsilon_{push}(L)$ は、下記の式 (4) で表される。下記の式 (3) 及び下記の式 (4) における「 j (矢印記号) i 」は、 x_j が x_i の近傍であることを意味する。

【 0 0 2 2 】

【数 3】

$$\varepsilon_{pull}(L) = \sum_{j \rightsquigarrow i} \|L(\vec{x}_i - \vec{x}_j)\|^2. \quad \dots (3)$$

30

【数 4】

$$\varepsilon_{push}(L) = \sum_{i, j \rightsquigarrow i} \sum_l (1 - y_{il}) \left[1 + \|L(\vec{x}_i - \vec{x}_j)\|^2 - \|L(\vec{x}_i - \vec{x}_l)\|^2 \right]_+. \quad \dots (4)$$

40

【 0 0 2 3 】

これら $\varepsilon_{pull}(L)$ 及び $\varepsilon_{push}(L)$ では、 k 近傍法に基づいてあるインスタンス i の近傍 k 個のインスタンス j が用いられる。例えば、上記の式 (3) に示す $\varepsilon_{pull}(L)$ によれば、同じラベルを持つインスタンス間の距離が大きい場合にペナルティが与えられる。さらに、上記の式 (4) に示す $\varepsilon_{push}(L)$ によれば、異なるラベルを持つインスタンス間の距離が小さい場合にペナルティが与えられる。

【 0 0 2 4 】

[3 . 2 特徴量選択]

特徴量の重要度を求める方法として、マハラノビス距離における M の対角成分を学習する方法が挙げられる。例えば、学習データや特徴量のスケール性に対応する側面から、マ

50

ハラノビス距離におけるMの対角成分以外の要素を0にすることで異なる特徴量間の関係性を無視し、各特徴量の重要度が学習される。

【0025】

この他、特徴量の重要度を求める方法として、決定木を用いる方法が知られている。例えば、決定木は、ノードを分割する特徴量の選択を繰り返すことにより生成される。このように特徴量を選択する場合、特徴量ごとに重要度が算出される。例えば、j番目の特徴量 f_j の重要度 $FI(f_j)$ は、下記の式(5)の通り、全てのノードにおける情報利得Iの総和を計算することにより算出することができる。

【0026】

【数5】

$$FI(f_i) = \sum_{t \in N_{f_j}}^n I(t, f_j) \quad \dots (5)$$

10

【0027】

上記の式(5)における「情報利得I」とは、親ノードから子ノードへ分割する際に得られる情報量を指す。ここで言う「親ノード」とは、特徴量の分岐で分割される前のノードを指し、また、「子ノード」とは、特徴量の分岐で分割された後のノードを指す。

【0028】

さらに、決定木における特徴量の分岐で分割される時に得られる情報量 $I(D_p, f)$ は、下記の式(6)によって表すことができる。下記の式(6)において、「f」は、分岐に選択された特徴量を指す。また、「 D_p 」は、親ノードを指す。さらに、「 D_{left} 」は、分岐後の左側の子ノードを指し、また、「 D_{right} 」は、分岐後の右側の子ノードを指す。さらに、「 N_p 」は、親ノード内のサンプルの個数を指す。さらに、「 N_{left} 」は、左側の子ノードのサンプルの個数を指し、また、「 D_{right} 」は、右側の子ノードのサンプルの個数を指す。

20

【0029】

【数6】

$$I(D_p, f) = I(D_p) - \frac{N_{left}}{N_p I(D_{left})} - \frac{N_{right}}{N_p I(D_{right})} \quad \dots (6)$$

30

【0030】

このような情報量を最大化する特徴量、言い換えれば親ノード及び子ノード間で不純度を最大限減らすことができる特徴量が決定木の分岐に選択される。あくまで一例として、上記の情報量として、ジニ係数やエントロピーなどを用いることができる。例えば、ジニ係数 $I_G(t)$ は、下記の式(7)によって算出することができる。また、エントロピー $I_H(t)$ は、下記の式(8)によって算出することができる。

【0031】

【数7】

$$I_G(t) = 1 - \sum_{i=1}^c p(i|t)^2 \quad \dots (7)$$

40

【数8】

$$I_H(t) = - \sum_{i=1}^c p(i|t) \log_2 p(i|t) \quad \dots (8)$$

【0032】

[3.3 マハラノビス距離学習の変形例]

50

上記のマハラノビス距離学習の変形例として、クラスタリング距離学習装置等が開示された文献が挙げられる。このクラスタリング距離学習装置では、上記の非特許文献3に記載のマハラノビス距離の行列Aの推定が学習データからの反復演算により行われる点が課題として挙げられている。

【0033】

このような課題を解決する側面から、クラスタリング距離学習装置では、マハラノビス距離の行列Aが収束するまで反復演算を行う非特許文献3に記載のアルゴリズムの代わりに次のような処理を行う。すなわち、クラスタリング距離学習装置では、非特許文献3に記載のマハラノビス距離の行列Aの推定アルゴリズムが、マハラノビス距離の行列Aの大きさを制限する拘束条件の下で行列Aとクラスタ中心 μ_k の関数Jを最小化する問題へ定式化される。このような定式化で得られる下記の式(9)に従って、マハラノビス距離の行列Aが計算される。

【0034】

【数9】

$$A = \frac{\sum_n \sum_k r_{nk} (x_n - \mu_k)(x_n - \mu_k)^T}{\left(\text{tr} \left(\sum_{n, kn', k'} r_{nk} r_{n'k'} (x_n - \mu_k)(x_n - \mu_k)^T (x_{n'} - \mu_{k'})(x_{n'} - \mu_{k'})^T \right) \right)^{1/2}} \quad \dots (7)$$

【0035】

上記の式(9)において、「 x_n 」及び「 $x_{n'}$ 」は、インスタンスの特徴量を指す。また、「 μ_k 」及び「 $\mu_{k'}$ 」は、クラスタk又はクラスタk'の中心を指す。また、「 r_{nk} 」及び「 $r_{n'k'}$ 」は、インスタンスとクラスタの対応関係を表す。例えば、「 r_{nk} 」は、インスタンス x_n がクラスタkに属する場合に“1”で表される一方でそれ以外の場合に“0”で表される。また、「 $r_{n'k'}$ 」は、インスタンス $x_{n'}$ がクラスタk'に属する場合に“1”で表される一方でそれ以外の場合に“0”で表される。また、「 $\text{tr}^T(AA)$ 」は、行列Aに対する拘束条件を指す。

【0036】

[3.4 ニューラルネットワークを用いる距離計量学習]

上記の距離計量学習は、上記のマハラノビス距離学習のように特徴量空間への線形な変換を学習する例に限定されず、クラス分類モデルの距離定義の部分にニューラルネットワークを適用することで、特徴量空間への非線形な変換を学習することもできる。

【0037】

[3.4.1 多クラス分類モデル]

モデル学習の結果として特徴量空間への非線形な変換を獲得できるという側面では、多クラス分類モデルの学習も、距離計量学習の一環とみなすこともできる。図2は、多クラス分類モデルの一例を示す図である。図2には、入力データのインスタンスが属するクラスのラベルを予測する多クラス分類モデル2の一例が示されている。このような多クラス分類モデル2に同一のラベルが付与された学習サンプルが入力される場合、i層目の隠れ層(中間層)から出力される隠れベクトル同士の距離が学習サンプル間で近くなるモデル学習が行われる。それ故、学習済みの多クラス分類モデル2が有する任意の隠れ層、例えば図2に示すn-1層目やn層目の隠れ層へ入力される隠れベクトルは、入力データが属するクラスのラベルに対応する距離空間上の位置に変換されているとみなすことができる。このことから、任意の隠れ層へ入力される隠れベクトルは、埋め込みベクトル(Embedded Vector)として用いることができる。

【0038】

[3.4.2 ディープメトリックラーニング]

ディープメトリックラーニングの一例として、Siamese Networkが知ら

10

20

30

40

50

れている。Siamese Networkでは、類似または非類似のペアに基づいて非線形かつ次元縮約を行いながら適切な距離空間上に入力データを写像する関数が学習される。

【0039】

図3は、Siamese Networkの構造の一例を示す図である。図3に示すように、Siamese Networkでは、類似または非類似のラベルが付与された2つのサンプルのペアが2つのニューラルネットワークNN1及びNN2へ入力される。このように2つのサンプルのペアが入力されるニューラルネットワークNN1及びNN2の間では、パラメータや層構造が共有される。そして、ニューラルネットワークNN1が出力する埋め込みベクトル及びニューラルネットワークNN2が出力する埋め込みベクトルから求まるサンプル間の距離が類似度として出力される。その上で、類似又は非類似のラベルに基づいて類似ペアの距離を近づける一方で、非類似ペアの距離を遠ざけるニューラルネットワークNN1及びNN2のパラメータが学習される。

10

【0040】

以下、距離空間への埋め込み、すなわち距離計量が行われるモデル、例えばSiamese NetworkのニューラルネットワークNN1やNN2等のことを「距離計量モデル」と記載する場合がある。

【0041】

[4. 課題の一側面]

例えば、学習済みのモデルを新たなタスクにも適合させる側面から、学習済みの距離計量モデルを新たなデータを用いて再学習するfine-tune等が行われる場合がある。この場合、これまでに挙げてきた距離計量モデルの再学習時に学習サンプルのデータ間に設定されるラベルが再学習後のモデルに悪影響を与える場合があるという側面がある。

20

【0042】

図4は、距離空間の一例を示す図である。図4には、再学習前の距離計量モデルにより埋め込まれる距離空間S1と、再学習後の距離計量モデルにより埋め込まれる距離空間S2とが左から順に並べて示されている。さらに、図4に示す距離空間S1及び距離空間S2には、再学習に用いられる学習データの各サンプルが再学習前または再学習後の埋め込みベクトルに基づいてクラスタリングされたクラスタリング結果が示されている。

【0043】

図4に示すように、再学習前の距離空間S1には、クラスタC1～クラスタC6の6つのクラスタが含まれる。ここで、あくまで一例として、上記6つのクラスタのうち、クラスタC1及びクラスタC2のペアに類似のラベルが設定された状態で再学習が行われた場合、距離空間S2への埋め込みが再学習される。すなわち、距離空間S2では、クラスタC1及びクラスタC2のペアに設定された類似のラベルに基づく再学習によってクラスタC1及びクラスタC2の間の距離が距離空間S1に比べて近づく。これに留まらず、距離空間S2では、クラスタC3及びクラスタC4の間の距離も距離空間S1に比べて近づく。これらクラスタC3及びクラスタC4の統合は、必ずしもモデルの設計者が意図するものとは限らず、想定外のモデルが再学習される場合がある。

30

【0044】

このような想定外の事態を避ける側面から、再学習時に全てのクラスタのペアに類似または非類似のラベルを設定するのはリソースの観点から必ずしも現実的とは言えず、ラベルの設定を最小限に留めたい側面がある。

40

【0045】

[5. 課題解決のアプローチの一側面]

そこで、本実施例では、再学習前および再学習過程の各時点で距離計量モデルが出力する埋め込みベクトルに基づいて互いが接近し、かつ類似のラベルが設定されていないクラスタのペアを出力する学習支援機能を提供する。このようなクラスタのペアが再学習過程で問い合わせられることによって、再学習時に設定される類似または非類似のラベルにより意図しないクラスタのペアが統合されるのを抑制し、もって再学習後の距離計量モデル

50

に悪影響を与えるのを抑制できる。

【 0 0 4 6 】

図 5 は、距離空間の一例を示す図である。図 5 では、距離計量学習のあくまで一例として、Siamese Network が採用される場合を例に挙げて説明を行う。図 5 には、再学習前の距離計量モデルにより埋め込まれる距離空間 S_{11} と、再学習過程の距離計量モデルにより埋め込まれる距離空間 S_{12} と、再学習後の距離計量モデルにより埋め込まれる距離空間 S_{13} とが左から順に並べて示されている。さらに、図 5 に示す距離空間 S_{11} 、距離空間 S_{12} 及び距離空間 S_{13} には、再学習に用いられる学習データの各サンプルが再学習前、再学習過程または再学習後の埋め込みベクトルに基づいてクラスタリングされたクラスタリング結果が示されている。

10

【 0 0 4 7 】

図 5 に示すように、再学習前の距離空間 S_{11} には、クラスタ $C_1 \sim$ クラスタ C_6 の 6 つのクラスタが含まれる。このような距離空間 S_{11} への埋め込みを行う再学習前の距離計量モデルを用いて、上記の学習支援機能は、上記 6 つのクラスタのうち、クラスタ C_1 及びクラスタ C_2 のペアに類似のラベルが設定された状態で、再学習を行う。この際、上記の学習支援機能は、所要エポック数、例えばロス関数の値が収束するエポック数やテストデータの正解率が一定値に達するエポック数よりも少ない再学習過程のエポック数、例えば 1 エポックの再学習を行う。このように 1 エポックの再学習が行われた結果、再学習過程では、距離空間 S_{12} への埋め込みが学習される。

【 0 0 4 8 】

ここで、上記の学習支援機能は、再学習前および再学習過程の各時点で距離計量モデルが出力する埋め込みベクトルに基づいて互いが接近し、かつ類似のラベルが設定されていないクラスタのペアを検出する。具体的には、クラスタ $C_1 \sim$ クラスタ C_6 ごとにクラスタ代表の埋め込みベクトル E_V が求められる。例えば、クラスタ代表の埋め込みベクトル E_V として、クラスタに属するサンプルを再学習前または再学習過程の距離計量モデルへ入力することにより得られる埋め込みベクトルの平均を用いることができる。

20

【 0 0 4 9 】

このようなクラスタ代表の埋め込みベクトルを用いて、上記の学習支援機能は、クラスタの移動方向を算出する。あくまで一例として、クラスタの移動方向は、再学習過程のクラスタ代表の埋め込みベクトルから再学習前のクラスタ代表の埋め込みベクトルを減算する計算により求めることができる。その上で、上記の学習支援機能は、あくまで一例として、下記の式 (10) にしたがって 2 つのクラスタの移動方向が略同一の直線上に存在するクラスタのペアを抽出する。例えば、下記の式 (10) において「 $\delta_{EV_cluster1}$ 」は、クラスタ C_1 の移動方向を指す。さらに、「 $\delta_{EV_cluster2}$ 」は、クラスタ C_2 の移動方向を指す。また、「 θ 」は、閾値を指す。このような式 (10) を満たすクラスタのペアが抽出される。

30

【 0 0 5 0 】

$$1 - |\cosine(\delta_{EV_cluster1}, \delta_{EV_cluster2})| < 1 \cdots (10)$$

【 0 0 5 1 】

さらに、上記の学習支援機能は、再学習前および再学習過程の各時点においてクラスタ間の距離を算出する。例えば、クラスタ間の距離は、クラスタのペアごとにクラスタ代表の埋め込みベクトルのユークリッド距離やコサイン距離を計算することにより求めることができる。その上で、上記の学習支援機能は、再学習前及び再学習過程の間でクラスタ間の距離の変化量を算出する。例えば、再学習前及び再学習過程の変化量は、再学習過程におけるクラスタ間の距離を再学習前のクラスタ間の距離で除算する計算により求めることができる。このように再学習前及び再学習過程の間で算出された変化量が所定の閾値、例えば θ 未満であるクラスタのペアが抽出される。

40

【 0 0 5 2 】

これら θ_1 及び θ_2 を用いて抽出されるクラスタのペアのうち類似のラベルが設定されていないクラスタのペアが再学習後に統合され得る問合せ対象として検出される。このよ

50

うなクラスタのペアがクライアント端末30等の任意の出力先に出力されることで、クラスタのペアの統合適否を受け付けることができる。

【0053】

図5に示す例で言えば、クラスタC1及びクラスタC2のペアとクラスタC3及びクラスタC4のペアが1及び2を用いて絞り込まれる。これらのうち、類似のラベルが設定されたクラスタC1及びクラスタC2のペアは、問合せ対象から除外される。その一方で、類似のラベルが設定されていないクラスタC3及びクラスタC4のペアは、問合せ対象として検出される。このようなクラスタC3及びクラスタC4のペアをクライアント端末30等に出力してクラスタのペアの統合適否、例えば再学習の中止やラベルの再設定などを受け付ける。例えば、再学習の中止のリクエストを受け付ける。これによって、モデルの設計者が意図しないクラスタの統合を抑制することができる。さらに、クラスタC3及びクラスタC4のペアに非類似のラベルが設定されると共に、クラスタC5及びクラスタC6のペアに類似のラベルが設定される。これによって、モデルの設計者が意図しないクラスタの統合を抑制した上で再学習を行う環境を調整することができる。

10

【0054】

このようなラベルの再設定後、上記の学習支援機能は、再設定されたラベルに基づいて所要エポック数の再学習を行う。例えば、クラスタC1及びクラスタC2のペアに類似のラベルが設定され、クラスタC3及びクラスタC4に非類似のラベルが設定されると共にクラスタC5及びクラスタC6に類似のラベルが設定された状態で再学習が行われる。この際、Siamese Network等の距離計量学習には、再学習前の距離計量モデルを用いることとしてもよいし、また、再学習過程の距離計量モデルを用いることとしてもかまわない。

20

【0055】

このように所要のエポック数の再学習が行われる結果、距離空間S13への埋め込みが再学習される。すなわち、距離空間S13では、クラスタC1及びクラスタC2の間の距離が距離空間S11及び距離空間S12に比べて近づくと共に、クラスタC5及びクラスタC6の間の距離が距離空間S11及び距離空間S12に比べて近づく。さらに、距離空間S13では、クラスタC3及びクラスタC4の間の距離が距離空間S11または距離空間S12に比べて接近するのを抑制したり、遠ざけたりすることができる。

【0056】

したがって、本実施例に係る学習支援機能によれば、モデルの設計者が意図しないクラスタのペアの統合が抑制されるので、再学習後の距離計量モデルが悪影響を受けるのを抑制することが可能になる。

30

【0057】

[6. サーバ装置10の構成]

図1には、実施例1に係るサーバ装置10の機能的構成の一例が示されている。図1に示すように、サーバ装置10は、通信インタフェース11と、記憶部13と、制御部15とを有する。なお、図1には、データの授受の関係を表す実線が示されているが、説明の便宜上、最小限の部分について示されているに過ぎない。すなわち、各処理部に関するデータの入出力は、図示の例に限定されず、図示以外のデータの入出力、例えば処理部及び処理部の間、処理部及びデータの間、並びに、処理部及び外部装置の間のデータの入出力が行われることとしてもかまわない。

40

【0058】

通信インタフェース11は、他の装置、例えばクライアント端末30との間で通信制御を行うインタフェースである。

【0059】

あくまで一例として、通信インタフェース11には、LANカードなどのネットワークインタフェースカードを採用することができる。例えば、通信インタフェース11は、クライアント端末30からラベルの設定、再学習の実行指示あるいは再学習の中止指示などを受け付ける。また、通信インタフェース11は、統合の問合せ対象のクラスタのペアをク

50

クライアント端末30へ送信したりする。

【0060】

記憶部13は、制御部15で実行されるOS(Operating System)を始め、各種プログラムに用いられるデータを記憶する機能部である。例えば、上記のプログラムには、上記の学習支援機能がモジュール化された学習支援プログラムを始め、上記の分類プログラムに学習支援プログラムがパッケージ化されたパッケージソフトウェアなどが対応し得る。

【0061】

あくまで一例として、記憶部13は、サーバ装置10における補助記憶装置に対応し得る。例えば、HDD(Hard Disk Drive)、光ディスクやSSD(Solid State Drive)などが補助記憶装置に対応する。この他、EPROM(Erasable Programmable Read Only Memory)などのフラッシュメモリも補助記憶装置に対応し得る。

10

【0062】

記憶部13は、制御部15で実行されるプログラムに用いられるデータの一例として、第1モデルデータ13M1と、学習データ14とを記憶する。これら第1モデルデータ13M1及び学習データ14以外にも、上記の学習支援プログラムによって参照されるデータ、例えば所要エポック数や再学習過程のエポック数などの再学習条件等が記憶部13に記憶されることとしてもよい。

【0063】

第1モデルデータ13M1は、再学習前の距離計量モデルのデータである。ここで言う「再学習前の距離計量モデル」は、あくまで一例として、再学習に用いられる学習データとは異なる学習データを用いて学習済みである距離計量モデルに対応し得る。例えば、第1モデルデータ13M1として、Siamese Networkを形成する入力層、隠れ層及び出力層の各層のニューロンやシナプスなどのモデルの層構造を始め、各層の重みやバイアスなどのモデルのパラメータが記憶部13に記憶される。

20

【0064】

学習データ14は、距離計量モデルの再学習に用いられるデータである。例えば、学習データ14には、文書データ14Aおよびラベルデータ14Bが含まれ得る。

【0065】

文書データ14Aは、文書のデータである。ここで言う「文書」は、距離計量モデルへ入力されるサンプルの一例に対応し得る。図6は、文書データ14Aの一例を示す図である。図6には、あくまで一例として、文書D1～文書D10の10個の文書が例示されている。図6に示す例で言えば、クラスタC1には、文書D1～D3が含まれる。さらに、クラスタC2には、文書D4～D6が含まれる。さらに、クラスタC3には、文書D7及びD8が含まれる。さらに、クラスタC4には、文書D9及びD10が含まれる。

30

【0066】

ここで、図6には、文書データ14Aの例として、テキストデータを示したが、テキストデータは、距離計量モデルへ入力する事前処理として、距離計量モデルへ入力可能な数値表現、例えばベクトル表現へ変換される。

【0067】

このような数値表現の変換には、あくまで一例として、Bag of words等を用いることができる。具体的には、文書D1～文書D10ごとに次のような処理が行われる。

40

【0068】

例えば、自然言語のテキストに形態素解析を適用することにより得られた文章の単語列のうち内容語の単語列が抽出される。図7は、単語列の抽出結果の一例を示す図である。図7には、図6に示された文書D1～文書D10ごとに単語列の抽出結果が示されている。例えば、図7に示すように、文書D1～文書D10の文章に対応する単語列から、機能語が除外される他、日付や日時等の特定の表現がストップワードとして除外された結果、内容語の単語列が抽出されている。

【0069】

このような文書D1～文書D10の内容語の単語列から、文書D1～文書D10の全て

50

の文書の辞書が生成される。例えば、「監視」、「APサーバ」、「DBサーバ」、「障害」、「エラー」、「発生」、「VEO000481436」及び「VEO000481437」などの単語を含む辞書が生成される。

【0070】

その上で、文書D1～文書D10ごとに、辞書の配列のうち単語列に含まれる単語の数値に「1」を割り当てる一方で、単語列に含まれない単語の数値に「0」を割り当てることで、Bag of wordsが生成される。

【0071】

例えば、文書D1及び文書D2を例に挙げれば、{監視：1，APサーバ：1，DBサーバ：0，障害：1，エラー：1，発生：1，VEO000481436：1，VEO000481437：0}といったベクトルが得られる。また、文書D3を例に挙げれば、{監視：1，APサーバ：1，DBサーバ：0，障害：0，エラー：1，発生：1，VEO000481436：1，VEO000481437：0}といったベクトルが得られる。さらに、文書D4及び文書D5を例に挙げれば、{監視：1，APサーバ：0，DBサーバ：1，障害：1，エラー：1，発生：1，VEO000481436：1，VEO000481437：0}といったベクトルが得られる。また、文書D6を例に挙げれば、{監視：1，APサーバ：0，DBサーバ：1，障害：0，エラー：1，発生：1，VEO000481436：1，VEO000481437：0}といったベクトルが得られる。さらに、文書D7及び文書D8を例に挙げれば、{監視：1，APサーバ：1，DBサーバ：0，障害：1，エラー：1，発生：1，VEO000481436：0，VEO000481437：1}といったベクトルが得られる。また、文書D9及び文書D10を例に挙げれば、{監視：1，APサーバ：0，DBサーバ：1，障害：1，エラー：1，発生：1，VEO000481436：0，VEO000481437：1}といったベクトルが得られる。

【0072】

このような事前処理により得られた文書D1～文書D10の8次元のベクトルが距離計量モデルへ入力され得る。以下、文書D1～文書D10のベクトルの各々を区別せずに総称できる場合、文書D1～文書D10のベクトルのことを「文書ベクトル」と記載する場合がある。

【0073】

ラベルデータ14Bは、クラスタのペアに設定されるラベルに関するデータである。

【0074】

例えば、ラベルデータ14Bは、クライアント端末30からラベルの設定を受け付けることにより生成され得る。あくまで一例として、図8に示すラベル設定画面200を介してラベルの設定を受け付けることができる。図8は、ラベル設定画面200の一例を示す図である。図8には、文書D1～文書D10のベクトルを再学習前の距離計量モデルへ入力することにより得られた埋め込みベクトルのクラスタリング結果がラベル設定画面200に表示されている。例えば、図8に示すように、ラベル設定画面200には、クラスタC1～C4が表示されている。このうち、クラスタC1には、文書D1～D3が含まれる。さらに、クラスタC2には、文書D4～D6が含まれる。さらに、クラスタC3には、文書D7及びD8が含まれる。さらに、クラスタC4には、文書D9及びD10が含まれる。これらクラスタC1～C4の表示と共に、ラベル設定画面200には、クラスタ内の文書間の埋め込みベクトルの距離が表示される。

【0075】

このようなラベル設定画面200上でクラスタのペアに類似のラベルを付与する操作を受け付ける。あくまで一例として、マウス等のポインティングデバイスが入力装置として用いられる場合、ドラッグ&ドロップの操作により文書のペアに類似のラベルを設定することができる。例えば、図8に示す例で言えば、ラベル設定画面200には、クラスタC1に属する文書D1をドラッグしてクラスタC2に属する文書D4へドロップすることで、クラスタC1及びクラスタC2のペアに類似のラベルが設定される例が示されている。

【0076】

10

20

30

40

50

このようにしてクラスタのペアに設定されるラベルがラベルデータ14Bとして保存される。なお、ここでは、指定が行われた文書のペアをクラスタのペアとみなしてラベルを設定する例を挙げたが、必ずしもクラスタのペアとみなさずともよく、文書のペアにラベルを設定することもできる。また、ここでは、ラベル設定の一例として、類似のラベルの設定を受け付ける例を挙げたが、当然のことながら非類似のラベルの設定を受け付けることもできる。さらに、ここでは、ユーザ操作によりラベルの設定を受け付ける例を挙げたが、ネットワークNWを介してラベルの設定を取得することもできれば、リムーバブルディスク等を含む内部または外部のストレージからラベルの設定を取得することとしてもかまわない。

【0077】

制御部15は、サーバ装置10の全体制御を行う機能部である。

【0078】

一実施形態として、制御部15は、CPU (Central Processing Unit) やMPU (Micro Processing Unit) などのハードウェアプロセッサにより実装することができる。ここでは、プロセッサの一例として、CPUやMPUを例示したが、汎用型および特化型を問わず、任意のプロセッサにより実装することができる。この他、制御部15は、ASIC (Application Specific Integrated Circuit) やFPGA (Field Programmable Gate Array) などのハードワイヤードロジックによって実現されることとしてもかまわない。

【0079】

制御部15は、上記の学習支援プログラムを実行することにより、図示しない主記憶装置として実装されるDRAM (Dynamic Random Access Memory) などのRAMのワークエリア上に図1に示す処理部を仮想的に実現する。なお、ここでは、上記の学習支援機能がモジュール化された学習支援プログラムが実行される例を挙げるが、サーバ装置10上で動作するプログラムはこれに限定されない。例えば、上記の分類プログラムに学習支援プログラムがパッケージ化されたパッケージソフトウェアが実行されることとしてもかまわない。

【0080】

例えば、制御部15は、図1に示すように、受付部15Aと、第1算出部15Bと、学習部15Cと、第2算出部15Dと、第3算出部15Eと、検出部15Fとを有する。

【0081】

受付部15Aは、再学習のリクエストを受け付ける処理部である。

【0082】

一実施形態として、受付部15Aは、クライアント端末30から図8に示されたラベル設定画面200に配置された再学習ボタン200Aの押下操作を受け付けることにより、fine-tune等の再学習のリクエストを受け付ける。そして、受付部15Aは、再学習のリクエストが受け付けられた場合、記憶部13から第1モデルデータ13M1及び学習データ14を読み出す。

【0083】

第1算出部15Bは、再学習前の埋め込みベクトルを算出する処理部である。

【0084】

一実施形態として、第1算出部15Bは、再学習前の距離計量モデルを用いて学習データ14の各クラスタの代表の埋め込みベクトルを算出する第1の距離計量処理を行う。例えば、第1算出部15Bは、記憶部13から読み出された第1モデルデータ13M1にしたがって図示しないメモリのワークエリア上に展開された再学習前の距離計量モデルへ学習データ14の各サンプルのベクトルを入力する。これによって、再学習前の距離計量モデルから埋め込みベクトルが出力される。

【0085】

より具体的には、クラスタC1~C4ごとにクラスタの代表とするサンプルである文書ベクトルが再学習前の距離計量モデルの入力層へ入力される。ここで、距離計量モデルの

10

20

30

40

50

入力層は文書 D 1 ~ 文書 D 1 0 の単語数「 8 」とし、出力層は 2 とする。さらに、学習データの各クラス C 1 ~ C 4 の代表とするサンプルの各々を文書 D 1、文書 D 4、文書 D 7、文書 D 9 とする。例えば、クラス C 1 の場合、文書 D 1 のベクトル { 監視 : 1 , A Pサーバ : 1 , D Bサーバ : 0 , 障害 : 1 , エラー : 1 , 発生 : 1 , VEO000481436 : 1 , VEO000481437 : 0 } が再学習前の距離計量モデルの入力層へ入力される。この結果、再学習前の距離計量モデルの出力層は、文書 D 1 の再学習前の埋め込みベクトル [- 5 , - 5] を出力する。同様に、文書 D 4 のベクトルが再学習前の距離計量モデルへ入力されることにより文書 D 4 の再学習前の埋め込みベクトル [- 5 , 5] が得られる。同様に、文書 D 7 のベクトルが再学習前の距離計量モデルへ入力されることにより文書 D 7 の再学習前の埋め込みベクトル [5 , 3] が得られる。同様に、文書 D 9 のベクトルが再学習前の距離計量モデルへ入力されることにより文書 D 4 の再学習前の埋め込みベクトル [5 , - 3] が得られる。

10

【 0 0 8 6 】

上記の第 1 の距離計量処理における `input` と `output` は、次の通りとなる。

【 0 0 8 7 】

`input` :

D1, D2 = { 監視 : 1 , A Pサーバ : 1 , D Bサーバ : 0 , 障害 : 1 , エラー : 1 , 発生 : 1 , VEO000481436 : 1 , VEO000481437 : 0 }

D3 = { 監視 : 1 , A Pサーバ : 1 , D Bサーバ : 0 , 障害 : 0 , エラー : 1 , 発生 : 1 , VEO000481436 : 1 , VEO000481437 : 0 }

20

D4, D5 = { 監視 : 1 , A Pサーバ : 0 , D Bサーバ : 1 , 障害 : 1 , エラー : 1 , 発生 : 1 , VEO000481436 : 1 , VEO000481437 : 0 }

D3 = { 監視 : 1 , A Pサーバ : 0 , D Bサーバ : 1 , 障害 : 0 , エラー : 1 , 発生 : 1 , VEO000481436 : 1 , VEO000481437 : 0 }

D7, D8 = { 監視 : 1 , A Pサーバ : 1 , D Bサーバ : 0 , 障害 : 1 , エラー : 1 , 発生 : 1 , VEO000481436 : 0 , VEO000481437 : 1 }

D9, D10 = { 監視 : 1 , A Pサーバ : 0 , D Bサーバ : 1 , 障害 : 1 , エラー : 1 , 発生 : 1 , VEO000481436 : 0 , VEO000481437 : 1 }

【 0 0 8 8 】

`output` : “ = ” は再学習前の距離計量モデルでの生成を意味する

30

D1_before_fine_tune = [-5, -5]

D4_before_fine_tune = [-5, 5]

D7_before_fine_tune = [5, 3]

D9_before_fine_tune = [5, -3]

【 0 0 8 9 】

図 9 は、埋め込みベクトルの一例を示す図である。図 9 には、再学習前の距離計量モデルにより埋め込まれるクラス C 1 ~ C 4 の代表とする文書 D 1、文書 D 4、文書 D 7 及び文書 D 9 の埋め込みベクトルがマッピングされている。図 9 に示すように、再学習前には、クラス C 1 の代表とする文書 D 1 が [- 5 , - 5] に埋め込まれると共に、クラス C 2 の代表とする文書 D 4 が [- 5 , 5] に埋め込まれる。さらに、クラス C 3 の代表とする文書 D 7 が [5 , 3] に埋め込まれると共に、クラス C 4 の代表とする文書 D 9 が [5 , - 3] に埋め込まれる。

40

【 0 0 9 0 】

学習部 1 5 C は、学習済みの距離計量モデルの再学習を行う処理部である。

【 0 0 9 1 】

一実施形態として、学習部 1 5 C は、学習データ 1 4 を用いて、第 1 モデルデータ 1 3 M 1 により定義された距離計量モデルの距離計量学習、すなわち再学習を行う。例えば、学習部 1 5 C は、文書 D 1 ~ D 1 0 のうち 2 つを組み合わせ得られる文書のペアごとに次のような処理を行う。すなわち、学習部 1 5 C は、文書のペアに設定された類似又は非類似のラベルに基づいて類似ペアの距離を近付ける一方で非類似ペアの距離を遠ざける S

50

iamese Networkのパラメータを再学習する。ここでは、あくまで一例として、学習部15Cは、クラスタC1及びクラスタC2のペアに設定された類似のラベルをSiamese Networkのパラメータの更新に用いることとする。このようにクラスタのペアにラベルが設定されている場合、各クラスタに属する全ての文書間の組合せに同一のラベル、すなわち類似のラベルが設定されているとみなして再学習が行われる。すなわち、上記の組合せには、文書D1及び文書D4のペア、文書D1及び文書D5のペア、文書D1及び文書D6のペアが含まれ得る。さらに、上記の組合せには、文書D2及び文書D4のペア、文書D2及び文書D5のペア、文書D2及び文書D6のペアが含まれ得る。さらに、上記の組合せには、文書D3及び文書D4のペア、文書D3及び文書D5のペア、文書D3及び文書D6のペアが含まれ得る。

10

【0092】

ここで、学習部15Cは、所要エポック数、例えばロス関数の値が収束するエポック数やテストデータの正解率が一定値に達するエポック数まで学習データ14を用いた再学習を繰り返すとは限らない。すなわち、学習部15Cは、所要エポック数よりも少ない再学習過程のエポック数、例えば1エポックの再学習を行う。

【0093】

上記の距離計量学習処理におけるinputとoutputは、次の通りとなる。

【0094】

input :

学習済みのSiamese Network

エポック回数：1回

学習データ :

類似ペア：(D1, D4), (D1, D5), (D1, D6), (D2, D4), (D2, D5), (D2, D6), (D3, D4), (D3, D5), (D3, D6)

20

【0095】

output :

再学習過程のSiamese Network

【0096】

このように学習部15Cの再学習により得られた再学習過程の距離計量モデルのパラメータ等が第2モデルデータ13M2として制御部15が参照するメモリのワークエリアに格納される。なお、ここでは、第2モデルデータ13M2が制御部15により参照されるメモリに格納される例を挙げたが、当然のことながら、任意のストレージ、例えば記憶部13が有する記憶領域に保存することとしてもかまわない。

30

【0097】

第2算出部15Dは、再学習過程の埋め込みベクトルを算出する処理部である。

【0098】

一実施形態として、第2算出部15Dは、再学習過程の距離計量モデルを用いて学習データ14の各クラスタの代表の埋め込みベクトルを算出する第2の距離計量処理を行う。例えば、第2算出部15Dは、上記の第2モデルデータ13M2にしたがって図示しないメモリのワークエリア上に展開された再学習過程の距離計量モデルへ学習データ14の各サンプルのベクトルを入力する。これによって、再学習過程の距離計量モデルから埋め込みベクトルが出力される。

40

【0099】

より具体的には、クラスタC1~C4ごとにクラスタの代表とするサンプルである文書ベクトルが再学習過程の距離計量モデルの入力層へ入力される。ここで、あくまで一例として、再学習前および再学習過程の間では、距離計量モデルのパラメータは異なれども層構造は共通することとする。例えば、クラスタC1の場合、文書D1のベクトル{監視：1, APサーバ：1, DBサーバ：0, 障害：1, エラー：1, 発生：1, VEO000481436：1, VEO000481437：0}が再学習過程の距離計量モデルの入力層へ入力される。この結果、再学習過程の距離計量モデルの出力層は、文書D1の再学習過程の埋め込み

50

ベクトル $[-5, -4]$ を出力する。同様に、文書 D 4 のベクトルが再学習過程の距離計量モデルへ入力されることにより文書 D 4 の再学習過程の埋め込みベクトル $[-5, 4]$ が得られる。同様に、文書 D 7 のベクトルが再学習過程の距離計量モデルへ入力されることにより文書 D 7 の再学習過程の埋め込みベクトル $[5, 2]$ が得られる。同様に、文書 D 9 のベクトルが再学習過程の距離計量モデルへ入力されることにより文書 D 4 の再学習過程の埋め込みベクトル $[5, -2]$ が得られる。

【0100】

上記の第 2 の距離計量処理における `input` と `output` は、次の通りとなる。

【0101】

`input` :

D1 = { 監視 : 1 , A Pサーバ : 1 , D Bサーバ : 0 , 障害 : 1 , エラー : 1 , 発生 : 1 , VEO000481436 : 1 , VEO000481437 : 0 }

D4 = { 監視 : 1 , A Pサーバ : 0 , D Bサーバ : 1 , 障害 : 1 , エラー : 1 , 発生 : 1 , VEO000481436 : 1 , VEO000481437 : 0 }

D7 = { 監視 : 1 , A Pサーバ : 1 , D Bサーバ : 0 , 障害 : 1 , エラー : 1 , 発生 : 1 , VEO000481436 : 0 , VEO000481437 : 1 }

D9 = { 監視 : 1 , A Pサーバ : 0 , D Bサーバ : 1 , 障害 : 1 , エラー : 1 , 発生 : 1 , VEO000481436 : 0 , VEO000481437 : 1 }

【0102】

`output` :

D1_process_of_fine_tune = $[-5, -4]$

D4_process_of_fine_tune = $[-5, 4]$

D7_process_of_fine_tune = $[5, 2]$

D9_process_of_fine_tune = $[5, -2]$

【0103】

図 10 は、埋め込みベクトルの一例を示す図である。図 10 には、再学習前の距離計量モデルにより埋め込まれるクラスタ C 1 ~ C 4 の代表とする文書 D 1、文書 D 4、文書 D 7 及び文書 D 9 の埋め込みベクトルが黒丸でマッピングされている。さらに、図 10 には、再学習過程の距離計量モデルにより埋め込まれるクラスタ C 1 ~ C 4 の代表とする文書 D 1、文書 D 4、文書 D 7 及び文書 D 9 の埋め込みベクトルが白丸でマッピングされている。図 10 に示すように、クラスタ C 1 の代表とする文書 D 1 は、再学習前には $[-5, -5]$ に埋め込まれている一方で、再学習過程では $[-5, -4]$ に埋め込まれている。さらに、クラスタ C 2 の代表とする文書 D 4 は、再学習前には $[-5, 5]$ に埋め込まれている一方で、再学習過程では $[-5, 4]$ に埋め込まれている。さらに、クラスタ C 3 の代表とする文書 D 7 は、再学習前には $[5, 3]$ に埋め込まれている一方で、再学習過程では $[5, 2]$ に埋め込まれている。さらに、クラスタ C 4 の代表とする文書 D 9 は、再学習前には $[5, -3]$ に埋め込まれている一方で、再学習過程では $[5, -2]$ に埋め込まれている。

【0104】

第 3 算出部 15 E は、再学習前および再学習過程の間でクラスタ間の移動パラメータを算出する処理部である。

【0105】

1 つの側面として、第 3 算出部 15 E は、再学習過程のクラスタ代表の埋め込みベクトルから再学習前のクラスタ代表の埋め込みベクトルを減算する計算により、クラスタの移動方向を算出する。このようにクラスタの移動方向を算出する場合における `input` と `output` は、次の通りとなる。

【0106】

`input` :

D1_before_fine_tune = $[-5, -5]$

D4_before_fine_tune = $[-5, 5]$

10

20

30

40

50

```

D7_before_fine_tune = [5, 3]
D9_before_fine_tune = [5, -3]
D1_process_of_fine_tune = [-5, -4]
D4_process_of_fine_tune = [-5, 4]
D7_process_of_fine_tune = [5, 2]
D9_process_of_fine_tune = [5, -2]

```

【 0 1 0 7 】

o u t p u t :

```

D1_delta = [0, 1]
D4_delta = [0, -1]
D7_delta = [0, -1]
D9_delta = [0, 1]

```

10

【 0 1 0 8 】

上記のクラスタの移動方向に基づいて、第3算出部15Eは、クラスタ間の進行角度の大きさを算出する。このようにクラスタ間の進行角度を算出する場合におけるinputとoutputは、次の通りとなる。

【 0 1 0 9 】

i n p u t :

```

D1_delta = [0, 1]
D4_delta = [0, -1]
D7_delta = [0, -1]
D9_delta = [0, 1]

```

20

【 0 1 1 0 】

o u t p u t :

```

クラスタC1代表の文書D1とクラスタC2代表の文書D4とのペアに関して
  1 - | C o s ( D1_delta , D4_delta ) | = 0
クラスタC1代表の文書D1とクラスタC3代表の文書D7とのペアに関して
  1 - | C o s ( D1_delta , D7_delta ) | = 0
クラスタC1代表の文書D1とクラスタC4代表の文書D9とのペアに関して
  1 - | C o s ( D1_delta , D9_delta ) | = 0
クラスタC2代表の文書D4とクラスタC3代表の文書D7とのペアに関して
  1 - | C o s ( D4_delta , D7_delta ) | = 0
クラスタC2代表の文書D4とクラスタC4代表の文書D9とのペアに関して
  1 - | C o s ( D4_delta , D9_delta ) | = 0
クラスタC3代表の文書D7とクラスタC4代表の文書D9とのペアに関して
  1 - | C o s ( D7_delta , D9_delta ) | = 0

```

30

【 0 1 1 1 】

他の側面として、第3算出部15Eは、再学習過程のクラスタ代表の埋め込みベクトル及び再学習前のクラスタ代表の埋め込みベクトルに基づいてクラスタ間の距離の変化量を算出する。あくまで一例として、再学習前及び再学習過程の変化量は、再学習過程におけるクラスタ間の距離を再学習前のクラスタ間の距離で除算する計算により求めることができる。このようにクラスタ間の距離の変化量を算出する場合におけるinputとoutputは、次の通りとなる。

40

【 0 1 1 2 】

i n p u t :

```

D1_before_fine_tune = [-5, -5]
D4_before_fine_tune = [-5, 5]
D7_before_fine_tune = [5, 3]
D9_before_fine_tune = [5, -3]
D1_process_of_fine_tune = [-5, -4]

```

50

D4_process_of_fine_tune = [-5, 4]

D7_process_of_fine_tune = [5, 2]

D9_process_of_fine_tune = [5, -2]

【0113】

output:

クラスタC1代表の文書D1とクラスタC2代表の文書D4とのペアに関して

$\text{after_distance}(D1, D4) / \text{before_distance}(D1, D4) = 8 / 10 = 0.8$

クラスタC1代表の文書D1とクラスタC3代表の文書D7とのペアに関して

$\text{after_distance}(D1, D7) / \text{before_distance}(D1, D7) = 136 / 164 = 0.829$

91

クラスタC1代表の文書D1とクラスタC4代表の文書D9とのペアに関して

$\text{after_distance}(D1, D9) / \text{before_distance}(D1, D9) = 104 / 104 = 1$

クラスタC2代表の文書D4とクラスタC3代表の文書D7とのペアに関して

$\text{after_distance}(D4, D7) / \text{before_distance}(D2, D7) = 104 / 104 = 1$

クラスタC2代表の文書D4とクラスタC4代表の文書D9とのペアに関して

$\text{after_distance}(D4, D9) / \text{before_distance}(D4, D9) = 136 / 164 = 0.829$

91

クラスタC3代表の文書D7とクラスタC4代表の文書D9とのペアに関して

$\text{after_distance}(D7, D9) / \text{before_distance}(D7, D9) = 4 / 6 = 0.67$

【0114】

上記の“before_distance”は、再学習過程におけるクラスタ間の距離を指し、また、上記の“after_distance”は、再学習過程におけるクラスタ間の距離を指す。あくまで一例を挙げれば、“before_distance(D1, D4)”は、再学習前におけるクラスタC1及びクラスタC2間の距離(=10)を指す。また、“after_distance(D1, D4)”は、図10に示すように、再学習過程におけるクラスタC1及びクラスタC2間の距離(=8)を指す。

10

20

【0115】

検出部15Fは、再学習後に統合され得るクラスタのペアを検出する処理部である。以下、再学習後に統合され得るクラスタのペアのことを「統合クラスタペア」と記載する場合がある。なお、検出部15Fは、出力部の一例に対応し得る。

【0116】

一実施形態として、検出部15Fは、第3算出部15Eにより算出されるクラスタ間の進行角度の大きさ及び第3算出部15Eにより算出されるクラスタ間の距離の変化量のうち少なくとも1つもしくは組合せに基づいて統合クラスタペアを検出する。

30

【0117】

あくまで一例として、検出部15Fは、クラスタ間の進行角度の大きさ及びクラスタ間の距離の変化量のAND条件で統合クラスタペアを検出することができる。例えば、検出部15Fは、クラスタ間の進行角度の大きさが所定の閾値1未満であり、かつクラスタ間の距離の変化量が所定の閾値2未満であるクラスタのペアを統合クラスタペアとして検出する。ここで、1を0.01とし、2を0.9としたとき、クラスタC1及びクラスタC2ペアと、クラスタC3及びクラスタC4のペアとが統合クラスタペアとして検出される。また、1を0.01とし、2を0.95としたとき、クラスタC1及びクラスタC2ペアと、クラスタC1及びクラスタC3ペアと、クラスタC2及びクラスタC4ペアと、クラスタC3及びクラスタC4のペアとが統合クラスタペアとして検出される。

40

【0118】

さらに、検出部15Fは、モデルの設計者の意図通りに統合されるクラスタのペアを除外する側面から、統合クラスタペアとして検出されたクラスタのペアのうち、類似のラベルが設定されたクラスタのペアを問合せの対象から除外する。例えば、図8に示すラベル設定画面200の例で言えば、ラベルデータ14Bには、クラスタC1及びクラスタC2のペアに類似のラベルが設定されているので、クラスタC1及びクラスタC2ペアが除外されてクラスタC3及びクラスタC4のペアが抽出される。

50

【 0 1 1 9 】

このように問合せ対象として検出された統合クラスタペアは、所定の出力先、例えばクライアント端末30に出力させることができる。図11は、問合せ画面の一例を示す図である。図11には、あくまで一例として、問合せ対象として検出された統合クラスタペアを含む問合せ画面400をラベル設定画面200の前面にポップアップ表示される例が示されている。図11に示すように、問合せ画面400には、再学習後にクラスタC3及びクラスタC4のペアが統合する可能性があることを警告するメッセージが表示される他、継続ボタン400Aと、中断ボタン400Bとが表示される。例えば、継続ボタン400Aに対する操作を受け付けた場合、再学習過程で再学習を中断せずに、所要のエポック数まで上記のラベルデータ14Bの設定のまま再学習が学習部15Cにより継続される。一方、中断ボタン400Bに対する操作を受け付けた場合、再学習過程で再学習を中断することができる。このような中断に加えて、ラベル設定画面上でラベルを再設定させることもできる。

10

【 0 1 2 0 】

図12は、ラベル設定画面の一例を示す図である。図12には、図11に示す問合せ画面400の中断ボタン400Bが操作された後に表示されるラベル設定画面210が示されている。図12に示すラベル設定画面210上で文書のペアを関連付ける操作を介して類似または非類似のラベルの再設定を受け付けることができる。例えば、図12に示すラベル設定画面210の例で言えば、クラスタC3に属する文書D7をドラッグしてクラスタC4に属する文書D9へドロップすることで、クラスタC3及びクラスタC4のペアに非類似のラベルが設定される例が示されている。このようにしてクラスタのペアに再設定されたラベルがラベルデータ14Bに上書き保存される。その後、学習ボタン210Aに対する操作を受け付けた場合、再設定が行われたラベルに基づいて学習部15Cに再学習を実行させることができる。なお、キャンセルボタン210Bに対する操作を受け付けた場合、再学習がキャンセルされる。

20

【 0 1 2 1 】

1つの側面として、類似のラベルが設定されていない統合クラスタペアを出力することで、次のような振る舞いを期待することができる。

【 0 1 2 2 】

すなわち、図8に示すラベル設定画面200では、クラスタC1及びクラスタC2を統合することを意図して、クラスタC1及びクラスタC2のペアに類似のラベルが設定される例が示されている。このようなラベル設定によって、文書D1～D3に出現する単語「APサーバ」および文書D4～D6に出現する単語「DBサーバ」のクラスタ形成における重要性が薄れる。これに伴って、“VE0000481436”や“VE0000481437”の重要性が距離空間への埋め込みを行う上で高まる。これらのことから、クラスタC1及びクラスタC2のペアへのラベル設定によってクラスタC1及びクラスタC2のペアの統合が発生する可能性が高まる。

30

【 0 1 2 3 】

このようにクラスタC1及びクラスタC2のペアの統合が発生する局面で類似のラベルが設定されていない統合クラスタペアが出力されることで、モデルの設計者が意図しないクラスタの統合の適否、例えば再学習の中止やラベルの再設定などの問合せを実現する。これによって、モデルの設計者が意図しないクラスタの統合を抑制することが可能である。さらに、クラスタC3及びクラスタC4のペアに対する非類似のラベルの設定等の再設定を受け付けることで、モデルの設計者が意図する通りの再学習を行うことが可能である。

40

【 0 1 2 4 】

なお、ここでは、クラスタC3及びクラスタC4のペアに対する非類似のラベルの設定が手動で受け付けられる例を挙げたが、類似のラベルが設定されていない統合クラスタペアに非類似のラベルを設定するレコメンドを自動的に出力することとしてもかまわない。

【 0 1 2 5 】

[7 . 処理の流れ]

50

図 1 3 は、実施例 1 に係る学習支援処理の手順を示すフローチャートである。この処理は、あくまで一例として、再学習のリクエストが受け付けられた場合等に開始される。図 4 に示すように、受付部 1 5 A は、記憶部 1 3 から第 1 モデルデータ 1 3 M 1 及び学習データ 1 4 を読み出す（ステップ S 1 0 1 ）。

【 0 1 2 6 】

続いて、第 1 算出部 1 5 B は、第 1 モデルデータ 1 3 M 1 に定義された再学習前の距離計量モデルを用いて学習データ 1 4 の各クラスタの代表の埋め込みベクトルを算出する第 1 の距離計量処理を行う（ステップ S 1 0 2 A ）。ステップ S 1 0 2 A で算出された再学習前のクラスタ代表の埋め込みベクトルは、第 1 算出部 1 5 B から第 3 算出部 1 5 E へ出力される（ステップ S 1 0 3 A ）。

10

【 0 1 2 7 】

上記のステップ S 1 0 2 A 及び上記のステップ S 1 0 3 A と並行して、下記のステップ S 1 0 2 B から下記のステップ S 1 0 5 B までの処理が実行される。

【 0 1 2 8 】

ステップ S 1 0 2 B では、学習部 1 5 C は、学習データ 1 4 を用いて、第 1 モデルデータ 1 3 M 1 により定義された距離計量モデルの距離計量学習、すなわち再学習を行う。ここで、学習データ 1 4 が再学習される回数は、所要エポック数よりも少ない再学習過程のエポック数が適用される。ステップ S 1 0 2 B の結果、再学習過程の距離計量モデルのパラメータ等が第 2 モデルデータ 1 3 M 2 として学習部 1 5 C から第 2 算出部 1 5 D へ出力される（ステップ S 1 0 3 B ）。

20

【 0 1 2 9 】

続いて、第 2 算出部 1 5 D は、第 2 モデルデータ 1 3 M 2 に定義された再学習過程の距離計量モデルを用いて学習データ 1 4 の各クラスタの代表の埋め込みベクトルを算出する第 2 の距離計量処理を行う（ステップ S 1 0 4 B ）。ステップ S 1 0 4 B で算出された再学習過程のクラスタ代表の埋め込みベクトルは、第 2 算出部 1 5 D から第 3 算出部 1 5 E へ出力される（ステップ S 1 0 5 B ）。

【 0 1 3 0 】

その後、第 3 算出部 1 5 E は、再学習前のクラスタ代表の埋め込みベクトル及び再学習過程のクラスタ代表の埋め込みベクトルに基づいてクラスタ間の移動パラメータ、例えばクラスタ間の進行角度の大きさやクラスタ間の距離の変化量を算出する（ステップ S 1 0 6 ）。ステップ S 1 0 6 で算出されたクラスタ間の移動パラメータは、第 3 算出部 1 5 E から検出部 1 5 F へ出力される（ステップ S 1 0 7 ）。

30

【 0 1 3 1 】

そして、検出部 1 5 F は、クラスタ間の進行角度の大きさ及びクラスタ間の距離の変化量のうち少なくとも 1 つもしくは組合せに基づいて再学習後に統合され得るクラスタのペアを検出する（ステップ S 1 0 8 ）。ステップ S 1 0 8 で検出されたクラスタのペアのうち類似のラベルが設定されていないクラスタのペアが所定の出力先、例えばクライアント端末 3 0 等へ出力される（ステップ S 1 0 9 ）。

【 0 1 3 2 】

[8 . 効果の一側面]

40

上述してきたように、本実施例に係るサーバ装置 1 0 は、再学習前および再学習過程の各時点で距離計量モデルが出力する埋め込みベクトルに基づいて互いが接近し、かつ類似のラベルが設定されていないクラスタのペアを検出する学習支援機能を提供する。したがって、本実施例に係るサーバ装置 1 0 によれば、モデルの設計者が意図しないクラスタのペアの統合が抑制されるので、再学習後の距離計量モデルが悪影響を受けるのを抑制することが可能である。

【実施例 2】

【 0 1 3 3 】

さて、これまで開示の装置に関する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。そこで、以下では、本発

50

明に含まれる他の実施例を説明する。

【0134】

[分散および統合]

また、図示した各装置の各構成要素は、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・統合の具体的形態は図示のものに限られず、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。例えば、受付部15A、第1算出部15B、学習部15C、第2算出部15D、第3算出部15Eまたは検出部15Fをサーバ装置10の外部装置としてネットワーク経由で接続するようにしてもよい。また、受付部15A、第1算出部15B、学習部15C、第2算出部15D、第3算出部15Eまたは検出部15Fを別の装置がそれぞれ有し、ネットワーク接続されて協働することで、上記のサーバ装置10の機能を実現するようにしてもよい。

10

【0135】

[学習支援プログラム]

また、上記の実施例で説明した各種の処理は、予め用意されたプログラムをパーソナルコンピュータやワークステーションなどのコンピュータで実行することによって実現することができる。そこで、以下では、図14を用いて、実施例1及び実施例2に係る学習支援プログラムを実行するコンピュータの一例について説明する。

【0136】

図14は、コンピュータのハードウェア構成例を示す図である。図14に示すように、コンピュータ100は、操作部110aと、スピーカ110bと、カメラ110cと、ディスプレイ120と、通信部130とを有する。さらに、このコンピュータ100は、CPU150と、ROM160と、HDD170と、RAM180とを有する。これら110~180の各部はバス140を介して接続される。

20

【0137】

HDD170には、図14に示すように、上記の実施例1で示した受付部15A、第1算出部15B、学習部15C、第2算出部15D、第3算出部15E及び検出部15Fと同様の機能を発揮する学習支援プログラム170aが記憶される。この学習支援プログラム170aは、図1に示した受付部15A、第1算出部15B、学習部15C、第2算出部15D、第3算出部15E及び検出部15Fの各構成要素と同様、統合又は分離してもかまわない。すなわち、HDD170には、必ずしも上記の実施例1で示した全てのデータが格納されずともよく、処理に用いるデータがHDD170に格納されればよい。

30

【0138】

このような環境の下、CPU150は、HDD170から学習支援プログラム170aを読み出した上でRAM180へ展開する。この結果、学習支援プログラム170aは、図14に示すように、学習支援プロセス180aとして機能する。この学習支援プロセス180aは、RAM180が有する記憶領域のうち学習支援プロセス180aに割り当てられた領域にHDD170から読み出した各種データを展開し、この展開した各種データを用いて各種の処理を実行する。例えば、学習支援プロセス180aが実行する処理の一例として、図13に示す処理などが含まれる。なお、CPU150では、必ずしも上記の実施例1で示した全ての処理部が動作せずともよく、実行対象とする処理に対応する処理部が仮想的に実現されればよい。

40

【0139】

なお、上記の学習支援プログラム170aは、必ずしも最初からHDD170やROM160に記憶されておらずともかまわない。例えば、コンピュータ100に挿入されるフレキシブルディスク、いわゆるFD、CD-ROM、DVDディスク、光磁気ディスク、ICカードなどの「可搬用の物理媒体」に各プログラムを記憶させる。そして、コンピュータ100がこれらの可搬用の物理媒体から各プログラムを取得して実行するようにしてもよい。また、公衆回線、インターネット、LAN、WANなどを介してコンピュータ100に接続される他のコンピュータまたはサーバ装置などに各プログラムを記憶させてお

50

き、コンピュータ100がこれらから各プログラムを取得して実行するようにしてもよい。

【符号の説明】

【0140】

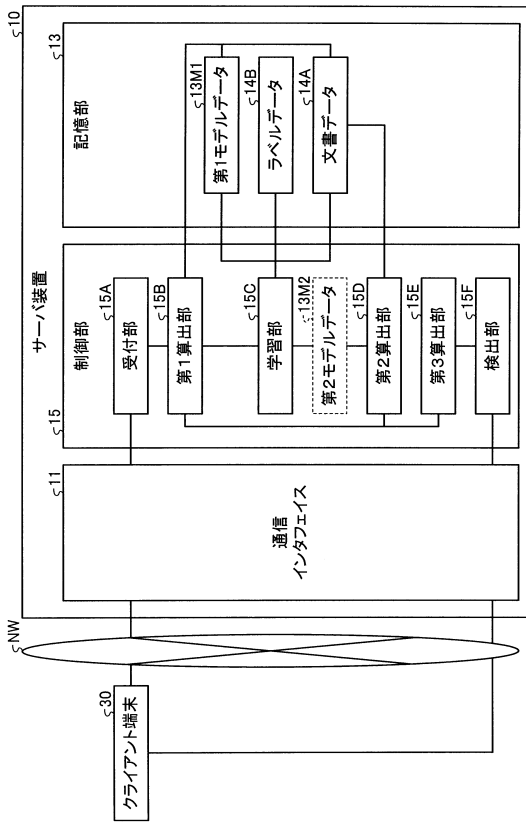
1	システム	
10	サーバ装置	
11	通信インタフェース	
13	記憶部	
13M1	第1モデルデータ	
14	学習データ	
14A	文書データ	10
14B	ラベルデータ	
15	制御部	
15A	受付部	
15B	第1算出部	
15C	学習部	
15D	第2算出部	
15E	第3算出部	
15F	検出部	
30	クライアント端末	20

30

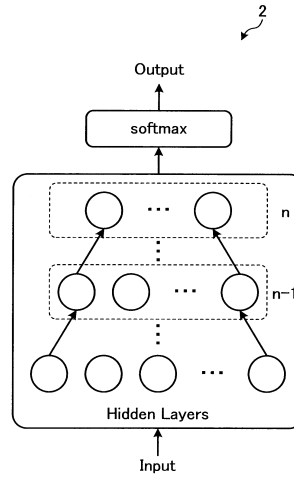
40

50

【図面】
【図 1】



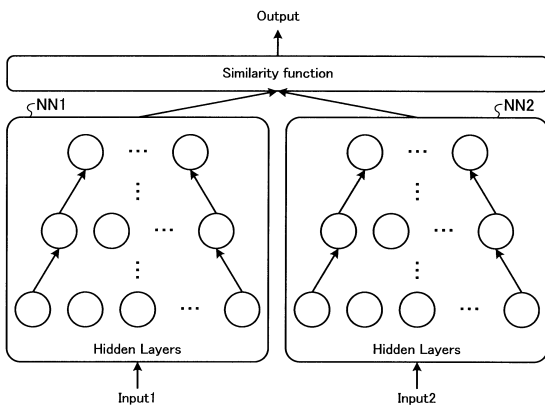
【図 2】



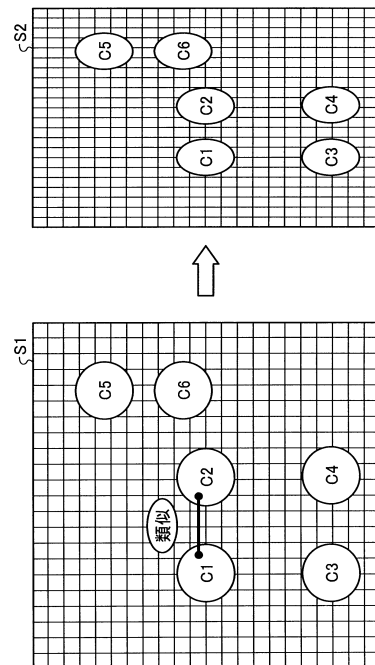
10

20

【図 3】



【図 4】

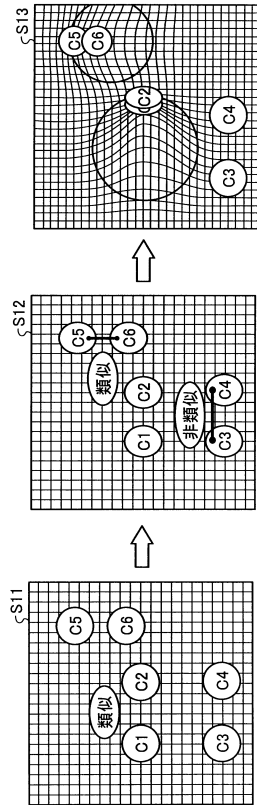


30

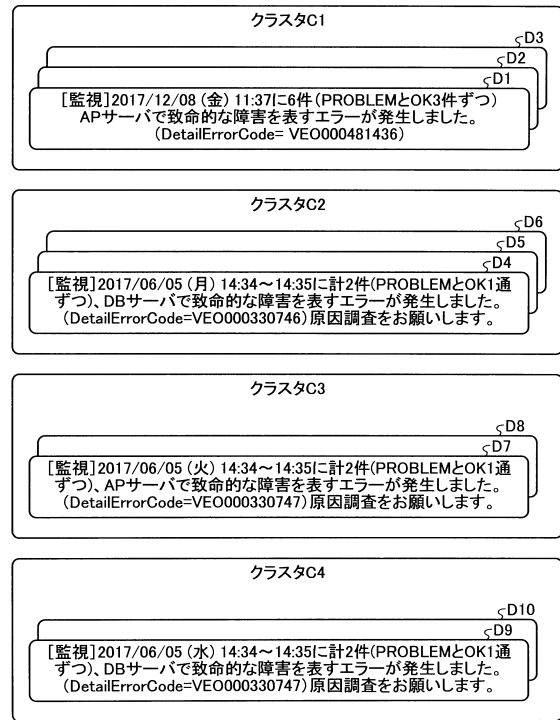
40

50

【 図 5 】



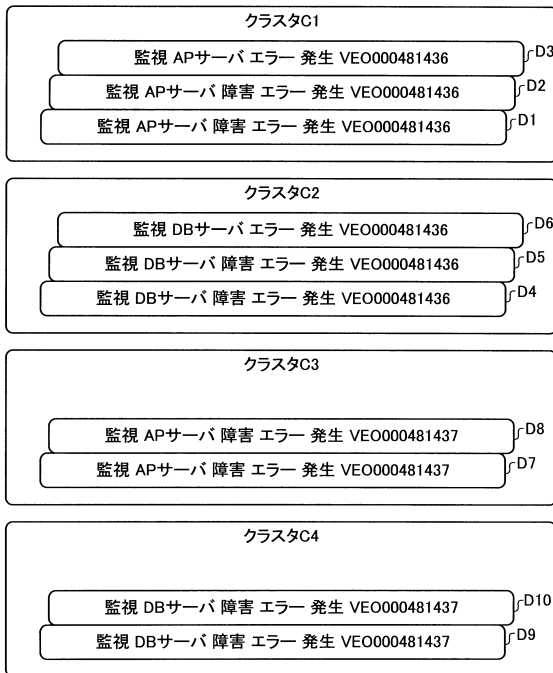
【 図 6 】



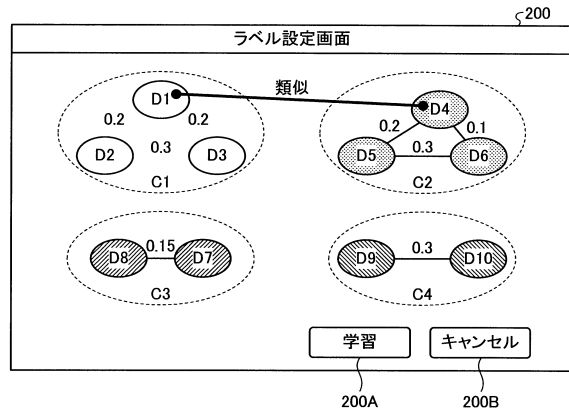
10

20

【 図 7 】



【 図 8 】

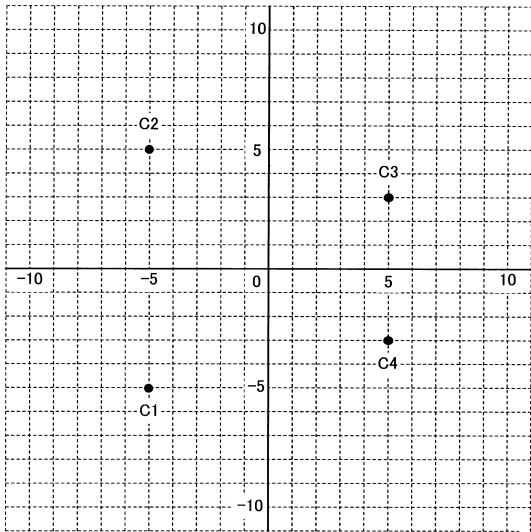


30

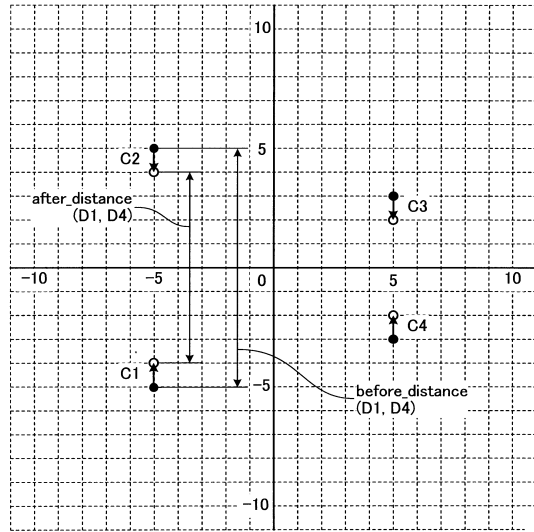
40

50

【 図 9 】

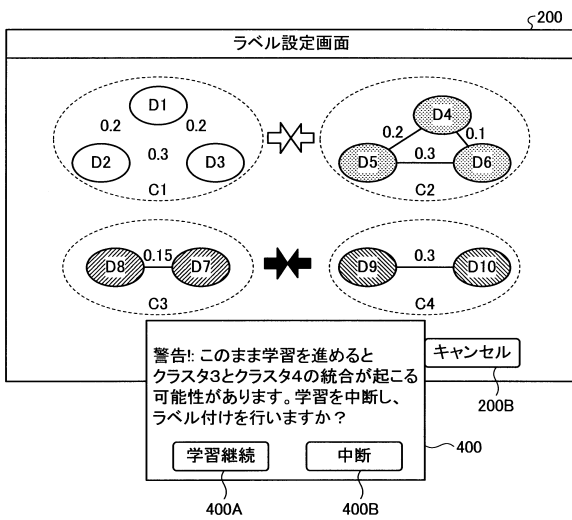


【 図 10 】

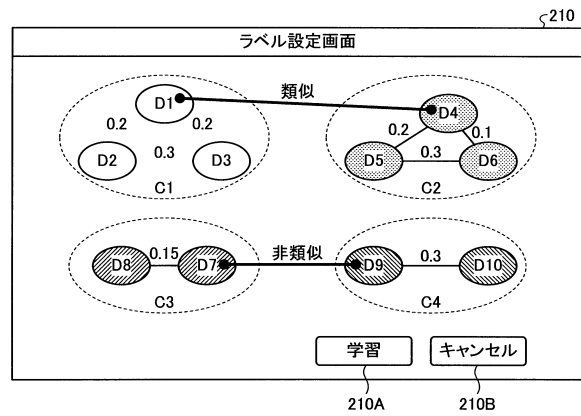


10

【 図 11 】



【 図 12 】



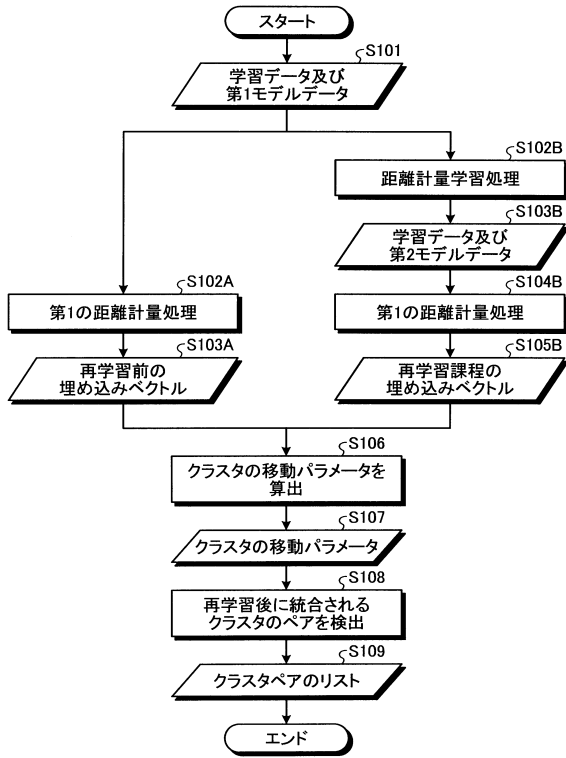
20

30

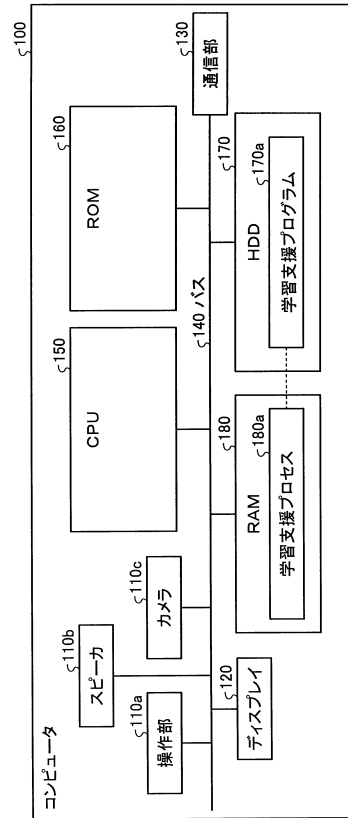
40

50

【図 13】



【図 14】



10

20

30

40

50

フロントページの続き

- (56)参考文献 特開 2 0 1 3 - 1 6 1 2 9 5 (J P , A)
岡田将吾(外 3 名), 「ユーザとの相互作用を通じた画像分類システムの開発」, 2010年度
人工知能学会全国大会(第24回)論文集, [CD-ROM], 2010年06月11日, Session ID: 2G2-OS9
-6, Pages 1-4, ISSN: 1347-9881, DOI: 10.11517/pjsai.JSAI2010.0_2G2OS96
- (58)調査した分野 (Int.Cl., D B 名)
G 0 6 N 2 0 / 0 0