

(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号
特許第7151604号
(P7151604)

(45)発行日 令和4年10月12日(2022.10.12)

(24)登録日 令和4年10月3日(2022.10.3)

(51)国際特許分類 F I
G 0 6 N 3/08 (2006.01) G 0 6 N 3/08 1 8 0

請求項の数 8 (全22頁)

(21)出願番号	特願2019-77274(P2019-77274)	(73)特許権者	000004226 日本電信電話株式会社 東京都千代田区大手町一丁目5番1号
(22)出願日	平成31年4月15日(2019.4.15)	(74)代理人	100107766 弁理士 伊東 忠重
(65)公開番号	特開2020-177290(P2020-177290 A)	(74)代理人	100070150 弁理士 伊東 忠彦
(43)公開日	令和2年10月29日(2020.10.29)	(74)代理人	100124844 弁理士 石原 隆治
審査請求日	令和3年7月26日(2021.7.26)	(72)発明者	田尻 兼悟 東京都千代田区大手町一丁目5番1号 日本電信電話株式会社内
		(72)発明者	渡辺 敬志郎 東京都千代田区大手町一丁目5番1号 日本電信電話株式会社内

最終頁に続く

(54)【発明の名称】 モデル学習装置、データ分析装置、モデル学習方法、及びプログラム

(57)【特許請求の範囲】

【請求項1】

訓練用データを用いて教師無の深層学習モデルの学習を行う学習手段と、
前記深層学習モデルにおける入力次元間の相関関係を算出する算出手段と、
相関関係が存在する次元の組毎に、前記訓練用データを用いて分析用モデルの学習を行
う分割モデル学習手段と
を備えるモデル学習装置。

【請求項2】

前記算出手段は、前記深層学習モデルにおける入力データの各次元の最終出力値に対す
る寄与度を算出し、当該寄与度に基づいて入力次元間の相関関係を算出する
請求項1に記載のモデル学習装置。

【請求項3】

請求項1又は2に記載の前記分割モデル学習手段により学習された分析用モデルを用い
てデータ分析を行うデータ分析手段を備えるデータ分析装置。

【請求項4】

訓練用データの次元を複数のグループに分割し、グループ毎に、分割された訓練用デー
タを用いて教師無の深層学習モデルの学習を行う部分学習手段と、
グループ毎に、前記深層学習モデルにおける入力次元間の相関関係を算出する算出手段
と、

グループ毎に、相関関係が存在する次元の組毎に前記訓練用データを用いて分割モデル

の学習を行う特徴量抽出手段と、

グループ毎の各分割モデルから得られる特徴量を用いて深層学習モデルを学習し、当該深層学習モデルにおける入力次元間の相関関係が存在する次元の組毎に、前記訓練用データを用いて分析用モデルの学習を行う学習手段と

を備えるモデル学習装置。

【請求項 5】

請求項 4 に記載の前記学習手段により学習された分析用モデルを用いてデータ分析を行うデータ分析手段を備えるデータ分析装置。

【請求項 6】

モデル学習装置が実行するモデル学習方法であって、
 訓練用データを用いて教師無の深層学習モデルの学習を行う学習ステップと、
 前記深層学習モデルにおける入力次元間の相関関係を算出する算出ステップと、
 相関関係が存在する次元の組毎に、前記訓練用データを用いて分析用モデルの学習を行う分割モデル学習ステップと
 を備えるモデル学習方法。

10

【請求項 7】

モデル学習装置が実行するモデル学習方法であって、
 訓練用データの次元を複数のグループに分割し、グループ毎に、分割された訓練用データを用いて教師無の深層学習モデルの学習を行う部分学習ステップと、
 グループ毎に、前記深層学習モデルにおける入力次元間の相関関係を算出する算出ステップと、

20

グループ毎に、相関関係が存在する次元の組毎に前記訓練用データを用いて分割モデルの学習を行うステップと、

グループ毎の各分割モデルから得られる特徴量を用いて深層学習モデルを学習し、当該深層学習モデルにおける入力次元間の相関関係が存在する次元の組毎に、前記訓練用データを用いて分析用モデルの学習を行うステップと

を備えるモデル学習方法。

【請求項 8】

コンピュータを、請求項 1、2 又は 4 に記載のモデル学習装置における各手段として機能させるためのプログラム。

30

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、深層学習を用いた分析に関するものであり、特に大規模のネットワーク機器から発生するログデータや I o T のセンサ群から得られる大量のデータに関する継続的な分析に関するものである。

【背景技術】

【0002】

分類問題（非特許文献 1）、未来予測（非特許文献 1）、異常検知（非特許文献 2）など様々なタスクに対し精度向上を目的として深層学習が用いられているが、深層学習の手法では訓練による深層学習モデルの構築と訓練済みモデルを用いた対象データの評価の 2 タームが存在し、これらのタームの中では入力されるデータの次元が等しくなければならないという前提が存在する。

40

【0003】

一方でネットワーク機器から発生するログデータや I o T のセンサ群から発生するデータは機器やセンサの交換や設定変更により、深層学習に入力されるデータ次元が変更される場合が存在し、この時訓練済みのモデルに対してデータの次元が変化したデータは入力できないためモデルの再訓練が必要となる。また、機械学習的な手法を用いる場合、分析対象データの種類（今回の問題設定の場合、ネットワーク機器やセンサの数に応じて増大）が多くなり過ぎると計算量が多くなりすぎる、また学習に必要なデータ量が増大する問

50

題がありスケールしない。

【先行技術文献】

【非特許文献】

【0004】

【文献】J. Schmidhuber, "Deep learning in neural networks: An overview", Neural Networks, 61, 2015.

R. Chalapathy and S. Chawla, "DEEP LEARNING FOR ANOMALY DETECTION: A SURVEY", arXiv:1901.03407, 2019.

G. E. Hinton and R. R. Salakhutdinov, "Reducing the Dimensionality of Data with Neural Networks" Science, 313 (5786), 2006. 10

X.Guo, X. Liu, E. Zhu and J. Yin, "Deep Clustering with Convolutional Autoencoders", ICONIP 2017.

P. Vincent, H. Larochelle, Y. Bengio and P. A. Manzagol "Extracting and composing robust features with denoising autoencoder", ICML 2008

D. P Kingma and M. Welling, "Auto-encoding variational Bayes", ICLR, 2014

S. Bach, A. Binder, G. Montavon, F. Klauschen, K. R. Muller and W. Samek, "On Pixel-Wise Explanations for Non-Linear Classifier Decisions by Layer-Wise Relevance Propagation", PloS one, 10(7), 2015.

A. Shrikumar, P. Greenside and A. Kundaje, "Learning important features through propagating activation differences", ICML, 2017 20

J. Macqueen, "Some methods for classification and analysis of multivariate observations", Proceedings of the fifth Berkeley symposium on mathematical statistics and probability, 1(14), 1967

【発明の概要】

【発明が解決しようとする課題】

【0005】

上記で述べたデータの次元変更によるモデルの再訓練が必要になった場合、訓練に必要なデータを収集する期間及びそのデータを用いてモデルを再訓練する期間では、上記で述べた分類、予測、異常検知等のタスクを実行できなくなる。また分析対象データの種類が多すぎる場合に分析用のモデルの学習が計算量及び学習データ量の観点から行えない場合がある。 30

【0006】

本発明は上記の点に鑑みてなされたものであり、モデルを用いてデータの分析を行う技術において、データの次元変更によるモデルの再訓練が必要になった場合でも、継続的に分析を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】

【0007】

開示の技術によれば、訓練用データを用いて教師無の深層学習モデルの学習を行う学習手段と、

前記深層学習モデルにおける入力次元間の相関関係を算出する算出手段と、 40

相関関係が存在する次元の組毎に、前記訓練用データを用いて分析用モデルの学習を行う分割モデル学習手段と

を備えるモデル学習装置が提供される。

【発明の効果】

【0008】

開示の技術によれば、モデルを用いてデータの分析を行う技術において、データの次元変更によるモデルの再訓練が必要になった場合でも、継続的に分析を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】

【0009】

【図 1】小規模データを取り扱う際の異常検知装置の機能ブロック図である。

【図 2】大規模データを取り扱う際の異常検知装置の機能ブロック図

【図 3】ネットワークログデータの前処理の例を説明するための図である。

【図 4】寄与度計算の概要を示す図である。

【図 5】分割モデル再学習の概要を示す図である。

【図 6】異常検知を例にした分析の例を示す図である。

【図 7】多段深層学習モデルによる相関関係取得の処理を示す図である。

【図 8】装置のハードウェア構成の例を示す図である。

【図 9】実施例 1、2 の処理を説明するためのフローチャートである。

【図 10】実施例 1 の処理を説明するためのフローチャートである。

10

【図 11】実施例 2 の処理を説明するためのフローチャートである。

【図 12】A E を用いた相関分割の例を示す図である。

【図 13】モデル分割と異常検知精度を示す図である。

【発明を実施するための形態】

【0010】

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。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。また、以下では、本発明を異常検知装置に適用する場合の例を示しているが、本発明は異常検知の分野に限らず、様々な分野に適用可能である。

【0011】

20

（実施の形態の概要）

データの次元変更が生じた場合でも、タスク実行不能時間をなくし継続的な分析を行うために、本実施の形態では、一つの深層学習モデルで入力データ全体を取り扱うのではなく、入力データの次元間の相関関係に基づき深層学習モデルを分割し、複数のモデルで入力データを取り扱う。この場合、入力次元が変更したとしても変化した次元が関与するモデルのみ再訓練を行い、その他無関係のモデルでは分析を継続することで、分析の継続性を確保する。

【0012】

また、分析対象データの種類数に関する分析用モデルの構築不可能性に関する問題に関しては、多段の相関関係取得モデルによるモデル分割により 1 つあたりのモデルが取り扱うデータの種類を削減することにより学習データ量や計算量を削減することを可能としている。

30

【0013】

以下では、具体的な実施の形態として、小規模データのモデル分割手法である実施例 1、及び、大規模データのモデル分割手法である実施例 2 を説明する。

【0014】

（実施例 1：小規模データのモデル分割手法）

まず、実施例 1 を説明する。

【0015】

<機能構成例>

40

実施例 1 における異常検知装置 100 の機能ブロックを図 1 に示す。

【0016】

図 1 に示すように、異常検知装置 100 は、データ収集部 110、データ前処理部 120、全体訓練部 130、深層学習モデル分割部 140、データ分析部 150 を有する。深層学習モデル分割部 140 は、寄与度計算部 141、相関関係算出部 142、分割モデル再学習部 143 を有する。各機能部の処理内容については後述する。

【0017】

なお、異常検知装置 100 は、モデル学習の機能を含むので、これをモデル学習装置と呼んでもよい。また、異常検知装置 100 は、データ分析の機能を含むので、これをデータ分析装置と呼んでもよい。

50

【 0 0 1 8 】

また、異常検知装置 1 0 0 からデータ分析部 1 5 0 を除いた装置をモデル学習装置と呼んでもよい。また、異常検知装置 1 0 0 からモデル学習用の機能部（全体訓練部 1 3 0 及び深層学習モデル分割部 1 4 0 ）を除いた装置をデータ分析装置と呼んでもよい。この場合のデータ分析装置には、分割モデル再学習部 1 4 3 で学習されたモデルが格納され、当該モデルがデータ分析に使用される。

【 0 0 1 9 】

< ハードウェア構成例 >

異常検知装置 1 0 0 は、例えば、コンピュータに、本実施の形態で説明する処理内容を記述したプログラムを実行させることにより実現可能である。例えば、モデルでデータを分析することは、データをコンピュータに入力し、モデルに相当するプログラムをコンピュータに実行させることで実現できる。

10

【 0 0 2 0 】

すなわち、異常検知装置 1 0 0 は、コンピュータに内蔵される CPU やメモリ等のハードウェア資源を用いて、当該異常検知装置 1 0 0 で実施される処理に対応するプログラムを実行することによって実現することが可能である。上記プログラムは、コンピュータが読み取り可能な記録媒体（可搬メモリ等）に記録して、保存したり、配布したりすることが可能である。また、上記プログラムをインターネットや電子メール等、ネットワークを通して提供することも可能である。

【 0 0 2 1 】

図 8 は、本実施の形態における上記コンピュータのハードウェア構成例を示す図である。図 8 のコンピュータは、それぞれバス B で相互に接続されているドライブ装置 1 0 0 0 、補助記憶装置 1 0 0 2 、メモリ装置 1 0 0 3 、CPU 1 0 0 4 、インタフェース装置 1 0 0 5 、表示装置 1 0 0 6 、及び入力装置 1 0 0 7 等を有する。

20

【 0 0 2 2 】

当該コンピュータでの処理を実現するプログラムは、例えば、CD-ROM 又はメモリカード等の記録媒体 1 0 0 1 によって提供される。プログラムを記憶した記録媒体 1 0 0 1 がドライブ装置 1 0 0 0 にセットされると、プログラムが記録媒体 1 0 0 1 からドライブ装置 1 0 0 0 を介して補助記憶装置 1 0 0 2 にインストールされる。但し、プログラムのインストールは必ずしも記録媒体 1 0 0 1 より行う必要はなく、ネットワークを介して他のコンピュータよりダウンロードするようにしてもよい。補助記憶装置 1 0 0 2 は、インストールされたプログラムを格納すると共に、必要なファイルやデータ等を格納する。

30

【 0 0 2 3 】

メモリ装置 1 0 0 3 は、プログラムの起動指示があった場合に、補助記憶装置 1 0 0 2 からプログラムを読み出して格納する。CPU 1 0 0 4 は、メモリ装置 1 0 0 3 に格納されたプログラムに従って、当該異常検知装置 1 0 0 に係る機能を実現する。インタフェース装置 1 0 0 5 は、ネットワークに接続するためのインタフェースとして用いられる。表示装置 1 0 0 6 はプログラムによる GUI (Graphical User Interface) 等を表示する。入力装置 1 0 0 7 はキーボード及びマウス、ボタン、又はタッチパネル等で構成され、様々な操作指示を入力させるために用いられる。

40

【 0 0 2 4 】

なお、前述したモデル学習装置、データ分析装置に関しても同様に、図 8 に示すようなコンピュータにプログラムを実行させることで実現できる。また、実施例 2 で説明する異常検知装置 2 0 0 （及びモデル学習装置、データ分析装置）に関しても同様に、図 8 に示すようなコンピュータにプログラムを実行させることで実現できる。

【 0 0 2 5 】

以下、異常検知装置 1 0 0 の各機能部の処理内容を詳細に説明する。

【 0 0 2 6 】

< データ収集部 1 1 0 、データ前処理部 1 2 0 >

データ収集部 1 1 0 は、本実施例で対象としている ICT システムのネットワークログ

50

データ（数値、テキスト）やIoTシステムのセンサデータを収集する。これらのデータはデータ前処理部120に送られて深層学習に用いることが可能な形状に整形される。

【0027】

図3に、ネットワークワークログデータに対する前処理により得られるデータの一例を示す。

【0028】

図3に示すように、整形されたデータは行列の形式をとり、横方向すなわち列のことをデータの次元と定義し、列の数を次元数、それぞれの列のことを各次元と呼ぶ。例えば、図3の例において、時刻00:00に得られた「メモリ（機器A）」の次元の値が0.2であることが示されている。

【0029】

<全体訓練部130>

モデル分割の対象とする分析手法は表現力の高い深層学習モデルであるので、モデル分割に用いる相関関係取得も深層学習モデルによって行う。そこで、全体訓練部130では整形されたデータを用いてデータ次元間の相関関係を調べるための深層学習モデルを構築する。相関関係を取得するための深層学習モデルとしては教師無のデータ特徴抽出モデルであるAutoEncoder(AE)（非特許文献3）、Convolutional AutoEncoder（非特許文献4）、Denoising AutoEncoder（非特許文献5）、Variational AutoEncoder(VAE)（非特許文献6）などを用いることができる。

【0030】

全体訓練部130が深層学習モデルを学習した後、学習に用いた整形済みデータと学習済み深層学習モデルを寄与度計算部141に入力する。

【0031】

<寄与度計算部141>

本実施例では、寄与度計算部141及び相関関係算出部142により、教師無の深層学習モデルに対し、出力側から入力側への逆伝播による深層学習モデルの解釈手法を活用し入力データの次元間の相関関係を算出する。より詳細には下記のとおりである。

【0032】

AE(VAE)及びその派生形は、深層学習モデル内部で入力データの特徴を抽出しつつ、出力が入力に近くなるように訓練を行うモデルである。ここでは、入力データの次元間相関関係の取得方法として深層学習モデルの解釈を目的として提案されている手法のうちデータの次元の重要度を計算する手法を用いる。このような手法としてLayer-wise relevance propagation(LRP)（非特許文献7）、DeepLIFT（非特許文献8）などが知られている。これらの手法は分類問題のテスト（分析）時にどの次元が分析結果に寄与していたのかを示す手法であるが、本実施例では、当該手法を訓練用データを復元するように学習するモデルであるAE(VAE)及びその派生形に応用する。

【0033】

寄与度計算部141が実行するLRPやDeepLIFTによる寄与度の算出方法について図4を参照しながら説明する。

【0034】

寄与度として(a)隣り合う層間毎の寄与度、及び(b)それらを繋げて最終出力値に対する入力の寄与度を求めることができる。以下ではLRP及びDeepLIFTによる解釈手法として提案されているもののうち一番簡単なものを例として説明する。ちなみに以降、寄与度算出部141の処理の説明における上付き文字は指数ではなく添え字である。

【0035】

(a)初めに層毎の寄与度について説明する。ここでは中間層(1層目)と中間層(2層目)を例にして説明する。なお、図面及び数式のイメージにおいて太字は多次元ベクトルを表す。明細書のテキストにおいては、多次元ベクトルを表す文字についてはその旨を記載する。

【0036】

10

20

30

40

50

(x^1, x^2) (x は多次元ベクトル)は図4の深層学習モデルの場合、1層目と2層目を結ぶ重み行列 W^1 、及びバイアス b^1 (b は多次元ベクトル)、非線形関数 f^1 を用いて、

【0037】

【数1】

$$x^2(x^1) = f^1(W^1x^1 + b^1) \quad (1)$$

10

と表わされる。 W 、 b 、 f は一般化されて、 k 層目と $k+1$ 層目に関しては W^k 、バイアス b^k 、非線形関数 f^k と表わされる。

【0038】

x^1 の j 次元目の x^2 の i 次元目に対する寄与度は

【0039】

【数2】

20

$$\text{LRP: } C_{ij}^1 = \left(\frac{\partial x_i^2}{\partial x_j^1}\right)x_j^1 \quad (2)$$

寄与度を比率にする場合：

【0040】

【数3】

30

$$C_{ij}^1 = \left(\frac{\partial x_i^2}{\partial x_j^1}\right)x_j^1 / \sum_j \left(\frac{\partial x_i^2}{\partial x_j^1}\right)x_j^1 \quad (3)$$

【0041】

【数4】

40

$$\text{DeepLIFT: } C_{ij}^1 = \frac{f^1(\sum_j W_{ij}^1 x_j^1 + b_i^1) - f^1(b_i^1)}{\sum_j W_{ij}^1 x_j^1} W_{ij}^1 x_j^1 \quad (4)$$

50

と表わされる。

【 0 0 4 2 】

寄与度計算部 1 4 1 において、訓練用データ全体もしくはサンプリングした一部の訓練データを訓練済みモデルの入力とし、各訓練データ毎に上記の寄与度を算出しその平均をとる。寄与度は 1 層目 2 層目間だけでなく 0 層目 1 層目間から、n 層目 n + 1 層目間まですべて計算を行う。図 4 の下側には、k = 0 ~ n、k 層目 m_k 次元、k + 1 層目 m_{k+1} 次元と仮定した場合における k 層目から k + 1 層目への寄与度行列 C^k が示されている。

【 0 0 4 3 】

(b) (a) で求めた寄与度 C_{ij}^k を元に、LRP 及び DeepLIFT それぞれで最終出力に対する入力データの各次元の寄与度を求めることができる。その結果は以下のようになる。以下の式において、i は出力値の次元、j は入力データの次元、 k_1 は 1 層目の次元である。

10

【 0 0 4 4 】

【数 5】

$$\text{LRP: } C_{ij} = \sum_{k_1 \dots k_n} \left(\frac{\partial \hat{x}_i}{\partial x_{k_n}^n} \right) \left(\frac{\partial x_{k_n}^n}{\partial x_{k_{n-1}}^{n-1}} \right) \dots \left(\frac{\partial x_{k_1}^1}{\partial x_j} \right) x_j \quad (5)$$

20

寄与度を比率にする場合：

【 0 0 4 5 】

【数 6】

$$C_{ij}^1 = \sum_{k_1 \dots k_n} \left(\frac{\partial \hat{x}_i}{\partial x_{k_n}^n} \right) \left(\frac{\partial x_{k_n}^n}{\partial x_{k_{n-1}}^{n-1}} \right) \dots \left(\frac{\partial x_{k_1}^1}{\partial x_j} \right) x_j / \sum_{j k_1 \dots k_n} \left(\frac{\partial \hat{x}_i}{\partial x_{k_n}^n} \right) \left(\frac{\partial x_{k_n}^n}{\partial x_{k_{n-1}}^{n-1}} \right) \dots \left(\frac{\partial x_{k_1}^1}{\partial x_j} \right) x_j \quad (6)$$

30

【 0 0 4 6 】

【数 7】

40

50

DeepLIFT:

$$C_{ij} = \sum_{k_1 \dots k_n} \frac{f^n(\sum_{k_n} W_{ik_n}^n x_{k_n}^n + b_i^n) - f^n(b_i^n)}{\sum_{k_n} W_{ik_n}^n x_{k_n}^n} \frac{f^{n-1}(\sum_{k_{n-1}} W_{k_n k_{n-1}}^{n-1} x_{k_{n-1}}^{n-1} + b_{k_n}^{n-1}) - f^{n-1}(b_{k_n}^{n-1})}{\sum_{k_{n-1}} W_{k_n k_{n-1}}^{n-1} x_{k_{n-1}}^{n-1}} \dots \frac{f^0(\sum_j W_{k_1 j}^0 x_j + b_{k_1}^1) - f^0(b_{k_1}^1)}{\sum_j W_{k_1 j}^0 x_j} W_{k_1 j}^0 x_j \quad (7)$$

この場合の寄与度も (a) の場合と同様、訓練用データ全体もしくはサンプリングした一部の訓練データを訓練済みモデルの入力とし、各訓練データ毎に上記の寄与度を算出しその平均をとる。

【 0 0 4 7 】

< 相関関係算出部 1 4 2 >

相関関係算出部 1 4 2 及び分割モデル再学習部 1 4 3 により、入力データの次元間の相関関係に基づき、入力データの次元をクラスタリングし、クラスタ毎に分析用の深層学習モデルを構築する。より具体的には下記のとおりである。

【 0 0 4 8 】

相関関係算出部 1 4 2 は寄与度計算部 1 4 1 で計算された寄与度を用いて入力次元の相関関係を取得する。相関関係取得法として、大まかに分けて、(1) 寄与度に対して閾値を設定する方法、(2) クラスタ数を設定する手法、の 2 種類の手法がある。以下、それぞれに対して詳しく説明する。

【 0 0 4 9 】

(1) 寄与度に対して閾値を設定する方法について、閾値を用いる段階を変えることで、下記の (イ)、(ロ) の 2 種類の手法がある。

【 0 0 5 0 】

(イ) 寄与度計算部 1 4 1 により前述した手法 (a) で計算された寄与度行列 C^k ($k = 0 \sim n$) を用いて以下のようなバイナリ行列 B^k ($k = 0 \sim n$) を作成する。

【 0 0 5 1 】

【 数 8 】

$$B_{ij}^k = \begin{cases} 0 & (C_{ij}^k < \varepsilon_k) \\ 1 & (C_{ij}^k \geq \varepsilon_k) \end{cases} \quad (8)$$

10

20

30

40

50

更に式(5)を用いて、
 【0052】
 【数9】

$$B = B^0 B^1 \dots B^n \quad (9)$$

を計算することで入出力の各次元が接続しているかいないかを表すバイナリ行列Bが得られる。AE及びVAEなどでは入力と出力の次元が等しいのでこの行列は正方行列であり、行方向が入力次元、列方向が出力次元となる。

【0053】

この正方行列を入出力次元数の列ベクトル B_i に分解し、以下の内積計算をすべての次元のペアで行い、相関関係の算出を行う。

【0054】

if $B_i \cdot B_j = 0$ 次元iと次元jに相関関係無し。

【0055】

if $B_i \cdot B_j > 0$ 次元iと次元jに相関関係あり。

すべての次元のペアに関して上の計算を行い、相関関係のあるグループ毎に次元をクラスタリングする。

【0056】

(ロ)寄与度計算部141により前述した手法(b)で計算された寄与度行列Cを列ベクトル C_i に分解し、各列ベクトルのペアワイズの距離を算出しそれに対して閾値を設定することで相関関係を算出する。ここで距離の定義として

L₁、L₂距離を含むミンコフスキー距離：

【0057】

【数10】

$$L_p = (\sum_k |x_k - y_k|^p)^{\frac{1}{p}} \quad (10)$$

コサイン類似度：

【0058】

【数11】

$$\cos(x, y) = \frac{\sum_k x_k y_k}{\sqrt{\sum_k x_k^2} \sqrt{\sum_k y_k^2}} \quad (11)$$

10

20

30

40

50

などを用いることができる。ただし、上記式の上付き文字は指数である。

【0059】

自身を含む全次元と相関関係のない次元が存在する場合、i) それらの次元をひとまとめにして1つの相関関係のグループとみなす場合やii) それらの次元は以降の分析に使用しないという2種類の取り扱いのいずれかを使用することができる。

【0060】

(2) クラスタ数を決めたくえでクラスタリングする手法としては、kMeans法(非特許文献9)を主に使用することができる。クラスタリングの入力として、(1)(ロ)と同様に C_i を用いる。この場合も孤立した次元が発生した場合i) それらを1つの相関関係グループとみなす場合とii) それぞれ独立な相関関係グループとみなす場合の2パターンのいずれかを使用することができる。

10

【0061】

<分割モデル再学習部143>

分割モデル再学習部143では、相関関係算出部142で得られた相関関係及び全体訓練部130で相関関係取得モデルを訓練するために用いた訓練用データを用いて相関関係が存在する次元毎に分析用モデルを用いて訓練を行う。この処理の具体例を図5に示す。

【0062】

相関関係算出部142により、図5に示す訓練用データが相関関係1{メモリ(機器A), CPU(機器A), ログ1出現数(正規化済み), ...}, 相関関係2{duration平均(source IP A), Bytes/min(機器A), ...}という2種類の相関関係に分割されたと仮定する。

20

【0063】

分割モデル再学習部143は、相関関係1に該当するデータを分析用モデル1に入力することで分析用モデル1の訓練を行い、相関関係2に該当するデータを分析用モデル2に入力することで分析用モデル2の訓練を行う。このように、相関関係毎のモデルにデータを入力することで訓練をやり直す。学習された各分析用モデルは、データ分析部150に格納される。

【0064】

<データ分析部150>

最後にデータ分析部150は、テスト(分析)に用いるデータを、分割モデル再学習部143で作成した複数のモデルに該当する次元毎に分けて入力し、分析結果を出力する。

30

【0065】

分析結果の出力の際に、最終的にすべてのモデルの出力をまとめて1つの結果として出力する必要がある場合には、その方法として、a) すべてのモデルから得られる出力結果をまとめて平均をとる、b) 各モデルの出力結果をバイナリ化してその平均値をとるなどの処理を行う。

【0066】

例えば、分類問題の場合、ある分析対象データの次元を分割し分割モデル再学習部143で得られた各モデルに次元分割したデータが入力された場合、それぞれのモデルは各ラベルに該当する確率を出力するが、一つの分析結果として出力する場合にはa) その確率を全モデルで平均し規格化する、b) 各モデルの確率を順位付けし投票制にするなどが考えられる。

40

【0067】

異常検知を例とした相関分割されたモデル群による分析の具体例を図6に示す。図6に示す例では、分析用データが、相関関係1に該当する次元のデータと、相関関係2に該当する次元のデータとに分割される。相関関係1に該当する次元のデータは、データ分析部150における分析用モデル1に入力され、相関関係2に該当する次元のデータは、データ分析部150における分析用モデル2に入力される。図6の例では、モデル1、モデル2のいずれから「異常」が出力され、これらをまとめて最終結果("異常")を出力する。

【0068】

50

解決しようとする課題で説明したデータの構造変化による継続不能性について、例えば、次元が減る場合は消滅した次元と相関を持つモデルを除いた他のモデルのみで分析を継続し、次元が増える場合はひとまず増えた次元を取り除いて分析を行いこれまでと挙動が大きく異なるように変化したモデルを次元変化の影響を受けたモデルとみなし、以降の分析ではそのモデルを取り除いた残りのモデルで分析を行う。

【0069】

(実施例2：大規模データのモデル分割手法)

次に実施例2を説明する。実施例2では、入力データの次元を恣意的に分割し、それぞれの中に存在する次元間の相関関係を実施例1で説明した手法で取得し、実施例1で説明した手法で相関に応じて分割された次元の組毎に特徴抽出のための教師無深層学習モデルを構築し、抽出された特徴量を利用して、全体的な相関関係を取得するという、深層学習モデルの段階的な使用により入力データの次元全体の相関関係を取得する。以下、より詳しく説明する。

10

【0070】

実施例1では、相関関係を取得するための深層学習モデルを訓練する全体訓練部130を導入している。しかし、取り扱うデータの次元が大きくなると全体訓練部130の学習が不能になる可能性がある。

【0071】

実施例1の方法で寄与度を計算しようとした場合に、次元数の大きさが原因で一つの相関学習モデルではデータ処理が行えずエラーが出るような場合、以下で説明する手法にて、大規模データに対する相関分割を行う。

20

【0072】

<機能構成例>

実施例2における異常検知装置200の機能ブロックを図2に示す。

【0073】

図2に示すように、異常検知装置200は、データ収集部210、データ前処理部220、部分訓練部230、部分深層学習モデル分割部240、全体訓練部250、全体深層学習モデル分割部260、データ分析部270を有する。部分深層学習モデル分割部240は、部分寄与度計算部241、部分相関関係算出部242、分割モデル特徴抽出部243を有する。全体深層学習モデル分割部260は、全体寄与度計算部261、全体相関関係算出部262、分割モデル再学習部263を有する。

30

【0074】

なお、異常検知装置200は、モデル学習の機能を含むので、これをモデル学習装置と呼んでもよい。また、異常検知装置200は、データ分析の機能を含むので、これをデータ分析装置と呼んでもよい。

【0075】

また、異常検知装置200からデータ分析部270を除いた装置をモデル学習装置と呼んでもよい。また、異常検知装置200からモデル学習用の機能部(部分訓練部230、部分深層学習モデル分割部240、全体訓練部250及び全体深層学習モデル分割部260)を除いた装置をデータ分析装置と呼んでもよい。この場合のデータ分析装置には、分割モデル再学習部263で学習されたモデルが入力され、当該モデルがデータ分析に使用される。

40

【0076】

また、実施例1の異常検知装置100(あるいは実施例1のモデル学習装置、データ分析装置)の機能と実施例2の異常検知装置200(あるいは実施例2のモデル学習装置、データ分析装置)の機能の両方を含む異常検知装置(あるいはモデル学習装置、データ分析装置)が用いられてもよい。当該異常検知装置(あるいはモデル学習装置)においては、例えば、入力データの規模が大きすぎて全体訓練部130での学習でエラーが生じた場合に、部分訓練部230での処理に移行する、といった処理を行うことができる。

【0077】

50

以下、実施例 2 の機能部の処理内容を説明する。

【0078】

<データ収集部 210、データ前処理部 220、部分訓練部 250>

データ収集部 210、データ前処理部 220 の処理内容は基本的には実施例 1 のデータ収集部 110、データ前処理部 120 の処理内容と同じである。

【0079】

実施例 2 では、データ前処理部 220 は、前処理において入力次元が大規模になった場合、入力次元を恣意的に分割する。なお、この分割は部分訓練部 230 が実施してもよい。分割の仕方としてネットワーク機器やセンサの実際の位置に基づく分割やデータの種類に基づく分割などを行うことができる。

10

【0080】

部分訓練部 230 は、恣意的に分割された訓練データを用いてその中での相関関係を取得するための深層学習モデルを作成する。ここで、部分訓練部 230 で用いる相関取得モデルは実施例 1 と同様 AE (VAE) 及びその派生形の教師無し深層学習モデルを使用することができる。実施例 2 では訓練用データが分割されているので、部分訓練部 230 は、分割された訓練データのそれぞれのモデルの訓練を行う。よって、複数のモデルが訓練される。

【0081】

<部分寄与度計算部 241、部分相関関係算出部 242>

部分寄与度計算部 141、部分相関関係算出部 242 の処理内容は、実施例 1 で説明した寄与度計算部 241、相関関係算出部 142 の処理内容と同じである。ただし、実施例 2 では、部分訓練部 230 により得られたそれぞれのモデルに対して、恣意的に分割された次元の中での相関関係を調べる。

20

【0082】

図 7 の (a) に具体例が示されている。図 7 の (a) に示す例では、前処理データを恣意的に 3 つのグループに分解したと仮定し、部分相関関係算出部 242 で、各グループについて相関関係を取得したことが示されている。図 7 の (a) において、グループ内の同一網掛けのノードが相関関係を持つことを示している。

【0083】

<分割モデル特徴抽出部 243>

分割モデル特徴抽出部 243 の処理内容は基本的には分割モデル再学習部 143 と同様である。分割モデル特徴抽出部 243 では、恣意的に分割された各グループの中の相関関係に基づいてモデルをさらに分割して、訓練用データを用いて AE あるいは VAE のような特徴を抽出するモデルを訓練させる。

30

【0084】

図 7 の (b) に具体例が示されている。分割モデル特徴抽出部 243 は、各グループの中で相関関係に基づいてモデル分割を行い、それぞれの相関関係の内部の特徴量を抽出するための深層学習モデルを学習する。図 7 の (b) は、グループ 1 が、分割モデル 1 ~ 3 等に分割され、それぞれで深層学習モデルの学習を行うことが示されている。3 グループの場合、グループ 2、3 も同様の学習を行う。

40

【0085】

<全体訓練部 250>

全体訓練部 250 では、分割モデル特徴抽出部 243 で訓練されたモデルに訓練データを入力した際に得られる次元が削減された中間層から出力されるデータを、全グループの全モデル分並べたものを相関取得モデルに対する入力とする。ここでも、実施例 1 の全体訓練部 130 と同様、全体訓練部 250 で用いる相関取得モデルとして AE (VAE) 及びその派生形を用いることができる。

【0086】

<全体寄与度計算部 261、全体相関関係算出部 262、分割モデル再学習部 263、データ分析部 270>

50

全体訓練部 2 5 0 で学習された深層学習モデルについても寄与度計算及び入力された中間層のデータに関する相関関係の算出をそれぞれ全体寄与度計算部 2 6 1 及び全体相関関係算出部 2 6 2 で行う。全体寄与度計算部 2 6 1 及び全体相関関係算出部 2 6 2 の処理内容自体は、実施例 1 の寄与度計算部 2 4 1 及び相関関係算出部 2 4 2 と同じである。

【 0 0 8 7 】

全体深層学習モデル分割部 2 6 0 において、入力に用いられた分割モデル特徴抽出部 2 4 3 のモデルの中間層がどの相関関係に属しているかは把握済みなので、これらの情報から入力次元全体の相関関係を把握することが可能になる。この相関関係に基づいて入力データの次元を分割し直し、実施例 1 と同様の分析モデル再学習及び、分析を分割モデル再学習部 2 6 3 及びデータ分析部 2 7 0 で行う。

10

【 0 0 8 8 】

例えば、恣意的に訓練データを 3 つのグループに分割した場合を考え、分割モデル特徴抽出部 2 4 3 により、グループ 1 について分割モデル 1 1、分割モデル 1 2、分割モデル 1 3 が得られ、グループ 2 について分割モデル 2 1、分割モデル 2 2 が得られ、グループ 3 について分割モデル 3 1、分割モデル 3 2、分割モデル 3 3、分割モデル 3 4 が得られたとする。

【 0 0 8 9 】

このとき、全体訓練部 2 5 0 では、分割モデル 1 1、分割モデル 1 2、分割モデル 1 3、分割モデル 2 1、分割モデル 2 2、分割モデル 3 1、分割モデル 3 2、分割モデル 3 3、及び分割モデル 3 4 のそれぞれの各中間層からの出力データが訓練を行う相関取得モデルに対する入力となる。仮に、各中間層の出力が 2 次元であるとすると。当該相関取得モデルの入力次元（及び出力次元）は、1 8 次元になる。

20

【 0 0 9 0 】

全体相関関係算出部 2 6 2 により、例えば、1 8 次元のうち、第 1 次元と第 1 0 次元の相関があることが分かったとする。そして、第 1 次元は分割モデル 1 1 に対応する相関関係に属し、第 1 0 次元は分割モデル 3 2 に対応する相関関係に属するとする。また、分割モデル 1 1 に対応する相関関係は、元の訓練データの第 2 次元、第 5 次元、第 6 次元であり、分割モデル 3 2 に対応する相関関係は、元の訓練データの第 4 次元、第 7 次元、第 8 次元であるとする。このとき、この相関に関して、分割モデル再学習部 2 6 2 では、元の訓練データの第 2 次元、第 4 次元、第 5 次元、第 6 次元、第 7 次元、第 8 次元の分析用モデルが学習されることになる。

30

【 0 0 9 1 】

図 7 の (c) に具体例が示されている。全グループの全モデルの中間層を横に並べて、全体訓練部 2 5 0、全体寄与度計算部 2 6 1、全体相関関係算出部 2 6 2 にて最終的に次元間相関関係を算出する。図 7 の (c) の例では、グループ 1、2、3 の同一網掛けの中間層が相関関係をもつ（出力次元が接合している）ため、グループ 1、2、3 に跨って同一網掛けの次元が相関関係を持っている。

【 0 0 9 2 】

（処理フロー）

図 9 ~ 図 1 1 のフローチャートを参照して、実施例 1、実施例 2 の全体の処理の流れを説明する。以下の説明の例では、最初に実施例 1 の異常検知装置 1 0 0 を使用し、データの規模に応じて、実施例 1 の異常検知装置 1 0 0 を継続使用、又は、実施例 2 の異常検知装置 2 0 0 を使用することとしている。なお、個々の機能部の処理内容については既に説明しているので、簡潔に説明している。

40

【 0 0 9 3 】

S 1 0 1 において、データ前処理部 1 2 0 により行列化されたデータを全体訓練部 1 3 0 に入力する。

【 0 0 9 4 】

S 1 0 2 において全体訓練部 1 3 0 が学習を行うが、データの規模が大きい場合、学習不能であることから（S 1 0 3 の No）、S 2 0 0（大規模データの相関分割（図 1 1）

50

）へ進む。それ以外の場合（S 1 0 3のYes）、S 1 0 4（小規模データの相関分割（図 1 0））へ進む。最初にS 1 0 4（小規模データの相関分割（図 1 0））へ進むほうを説明する。

【 0 0 9 5 】

図 1 0 の S 1 0 5 において、寄与度計算部 1 4 1 は寄与度を計算する。S 1 0 6 において、相関関係算出部 1 4 2 は、相関関係を算出し、訓練データの次元を分割する。

【 0 0 9 6 】

S 1 0 7 において、分割モデル再学習部 1 4 3 は、分割された次元毎に分析用モデルを学習する。S 1 0 8 において、データ分析部 1 5 0 は、分割モデル再学習部 1 4 3 により学習された分析用モデルを用いて、テストデータに対する分析を実行する。

10

【 0 0 9 7 】

次に、S 2 0 0（大規模データの相関分割（図 1 1））へ進むほうを説明する。

【 0 0 9 8 】

S 2 0 1 において、データ前処理部 2 2 0 が、前処理された行列データの次元を恣意的にいくつか分割する。S 2 0 2 において、部分訓練部 2 3 0 は、各分割データを用いて、分割されたグループ毎のモデルの学習を行う。

【 0 0 9 9 】

S 2 0 3 において、部分寄与度計算部 2 4 1 は、モデル毎に寄与度計算を実施する。S 2 0 4 において、部分相関関係算出部 2 4 2 は、モデル毎に相関関係を算出し、モデル毎に次元の分割を行う。S 2 0 5 において、分割モデル特徴抽出部 2 4 3 は、分割されたモデル毎にモデル再学習を行う。

20

【 0 1 0 0 】

S 2 0 6 において、全体訓練部 2 5 0 は、分割モデル特徴抽出部 2 4 3 で得られた特徴量を用いてモデル学習を行う。S 2 0 7 において、全体寄与度計算部 2 6 1 が寄与度計算を行う。S 2 0 8 において、全体相関関係算出部 2 6 2 が、相関関係を算出し、相関関係に基づいて次元の分割を行う。

【 0 1 0 1 】

S 2 0 9 において、分割モデル再学習部 2 6 3 は、相関に基づき分割されたモデルの再学習を行う。S 2 1 0 において、データ分析部 2 7 0 は、分割モデル再学習部 2 6 3 により学習された分析用モデルを用いて、テストデータに対する分析を実行する。

30

【 0 1 0 2 】

（実施の形態に係る技術の効果について）

実施例 1、2 を用いて説明した本実施の形態に係る技術により、データの相関関係という特性に基づいてモデルを分割することで、分析の精度を落とすことなくデータ構造変化が発生した場合の分析タスクの継続不能性という問題に対応することが可能となる。

【 0 1 0 3 】

以下では異常検知のタスクを例に挙げ、その精度を落とすことなくモデルの分割が可能であることを示す。

【 0 1 0 4 】

K S L _ K D D というネットワーク侵入検知系のベンチマークデータに対して A E を用いた異常検知を相関関係に基づいて分割した結果を示す。図 1 2 は A E を用いた相関分割の結果（一部）を示しており、各 は下側から入力、中間層、出力の各次元を表しており、入力の次元の で同じ網掛けのものが本手法により相関関係があると判定されたものである。

40

【 0 1 0 5 】

また相関関係の取得方法としては各層間のリンクに対する閾値決定によってリンクの切断を行いその上で出力同士が繋がっている場合に相関があるとみなしている。さらにこの相関関係に基づき異常検知のための深層学習モデルを分割し異常検知を行った結果を図 1 3 に示す。

【 0 1 0 6 】

50

図 1 3 における A U C が異常検知の精度を表しており、A U C が高い方が異常検知の精度が高い。図 1 3 は相関関係を決定するための閾値によって異常検知精度が変わること、及びモデル分割を行うとモデル分割を行わない場合（閾値 = 0）よりも精度がよくなる場合があるということを示しており、モデル分割をタスクの精度を落とさず実行できることからこの手法の有用性を表している。

【 0 1 0 7 】

（実施の形態のまとめ）

本実施の形態により、少なくとも下記の各項に記載されたモデル学習装置、データ分析装置、モデル学習方法、及びプログラムが提供される。

（第 1 項）

訓練用データを用いて教師無の深層学習モデルの学習を行う学習手段と、
前記深層学習モデルにおける入力次元間の相関関係を算出する算出手段と、
相関関係が存在する次元の組毎に、前記訓練用データを用いて分析用モデルの学習を行う分割モデル学習手段と
を備えるモデル学習装置。

10

（第 2 項）

前記算出手段は、前記深層学習モデルにおける入力データの各次元の最終出力値に対する寄与度を算出し、当該寄与度に基づいて入力次元間の相関関係を算出する
第 1 項に記載のモデル学習装置。

（第 3 項）

第 1 項又は第 2 項に記載の前記分割モデル学習手段により学習された分析用モデルを用いてデータ分析を行うデータ分析手段を備えるデータ分析装置。

20

（第 4 項）

訓練用データの次元を複数のグループに分割し、グループ毎に、分割された訓練用データを用いて教師無の深層学習モデルの学習を行う部分学習手段と、
グループ毎に、前記深層学習モデルにおける入力次元間の相関関係を算出する算出手段と、
グループ毎に、相関関係が存在する次元の組毎に前記訓練用データを用いて分割モデルの学習を行う特徴量抽出手段と、

グループ毎の各分割モデルから得られる特徴量を用いて深層学習モデルを学習し、当該深層学習モデルにおける入力次元間の相関関係が存在する次元の組毎に、前記訓練用データを用いて分析用モデルの学習を行う学習手段と
を備えるモデル学習装置。

30

（第 5 項）

第 4 項に記載の前記学習手段により学習された分析用モデルを用いてデータ分析を行うデータ分析手段を備えるデータ分析装置。

（第 6 項）

モデル学習装置が実行するモデル学習方法であって、
訓練用データを用いて教師無の深層学習モデルの学習を行う学習ステップと、
前記深層学習モデルにおける入力次元間の相関関係を算出する算出ステップと、
相関関係が存在する次元の組毎に、前記訓練用データを用いて分析用モデルの学習を行う分割モデル学習ステップと
を備えるモデル学習方法。

40

（第 7 項）

モデル学習装置が実行するモデル学習方法であって、
訓練用データの次元を複数のグループに分割し、グループ毎に、分割された訓練用データを用いて教師無の深層学習モデルの学習を行う部分学習ステップと、
グループ毎に、前記深層学習モデルにおける入力次元間の相関関係を算出する算出ステップと、

グループ毎に、相関関係が存在する次元の組毎に前記訓練用データを用いて分割モデル

50

の学習を行うステップと、

グループ毎の各分割モデルから得られる特徴量を用いて深層学習モデルを学習し、当該深層学習モデルにおける入力次元間の相関関係が存在する次元の組毎に、前記訓練用データを用いて分析用モデルの学習を行うステップと

を備えるモデル学習方法。

(第8項)

コンピュータを、第1項、第2項又は第4項に記載のモデル学習装置における各手段として機能させるためのプログラム。

【0108】

以上、本実施の形態について説明したが、本発明はかかる特定の実施形態に限定されるものではなく、特許請求の範囲に記載された本発明の要旨の範囲内において、種々の変形・変更が可能である。

10

【符号の説明】

【0109】

- 100 異常検知装置
- 110 データ収集部
- 120 データ前処理部
- 130 全体訓練部
- 140 深層学習モデル分割部
- 141 寄与度計算部
- 142 相関関係算出部
- 143 分割モデル再学習部
- 150 データ分析部
- 200 異常検知装置
- 210 データ収集部
- 220 データ前処理部
- 230 部分訓練部
- 240 部分深層学習モデル分割部
- 241 部分寄与度計算部
- 242 部分相関関係算出部
- 243 分割モデル特徴抽出部
- 250 全体訓練部
- 260 全体深層学習モデル分割部
- 261 全体寄与度計算部
- 262 全体相関関係算出部
- 263 分割モデル再学習部
- 270 データ分析部
- 1000 ドライブ装置
- 1002 補助記憶装置
- 1003 メモリ装置
- 1004 CPU
- 1005 インタフェース装置
- 1006 表示装置
- 1007 入力装置

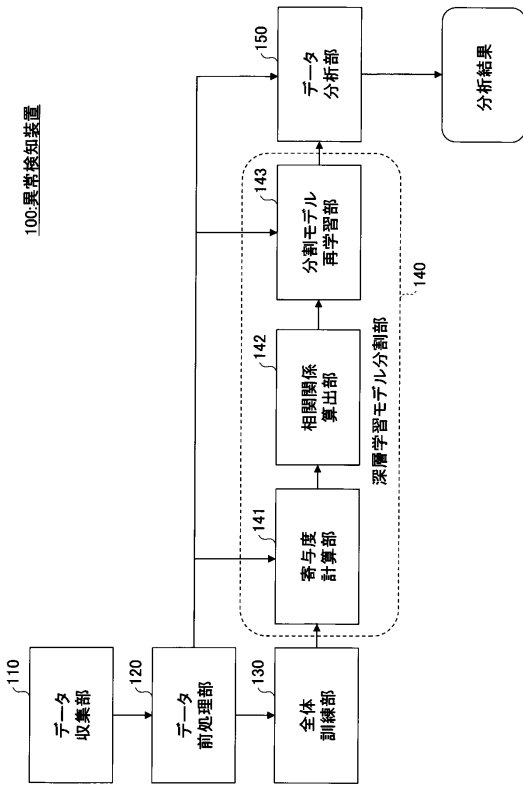
20

30

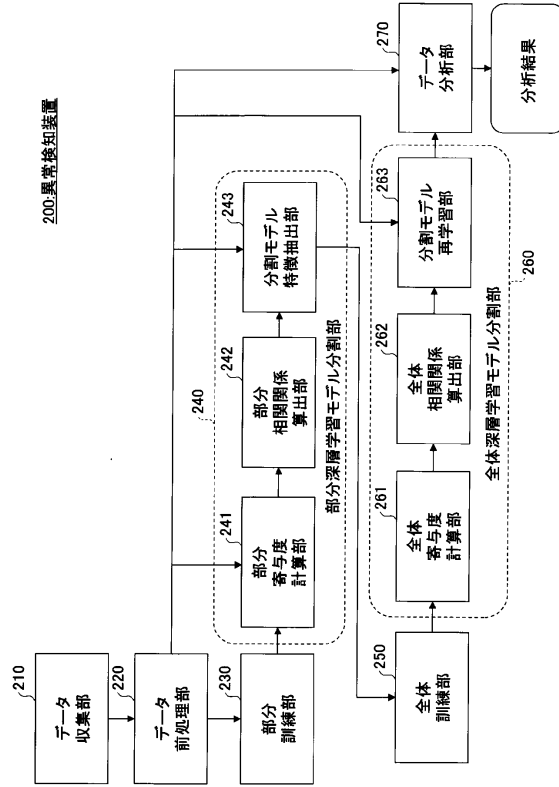
40

50

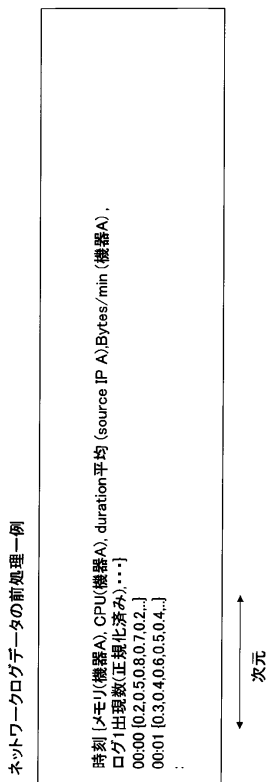
【図1】
【図面】



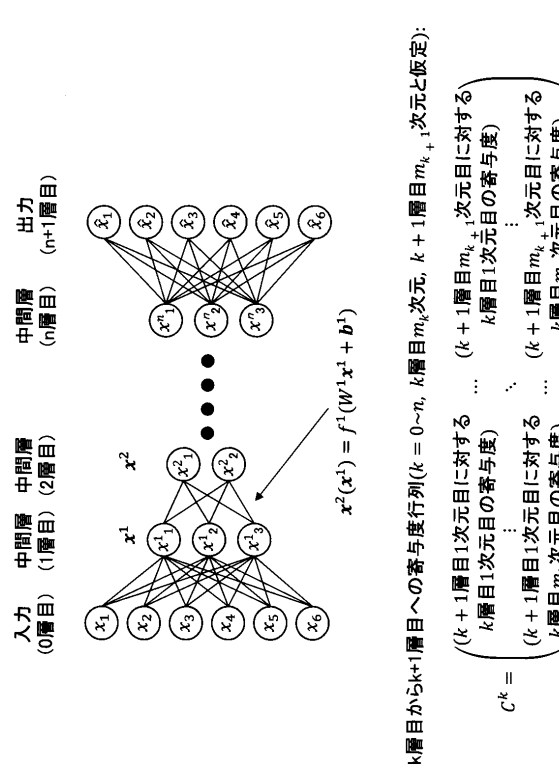
【図2】



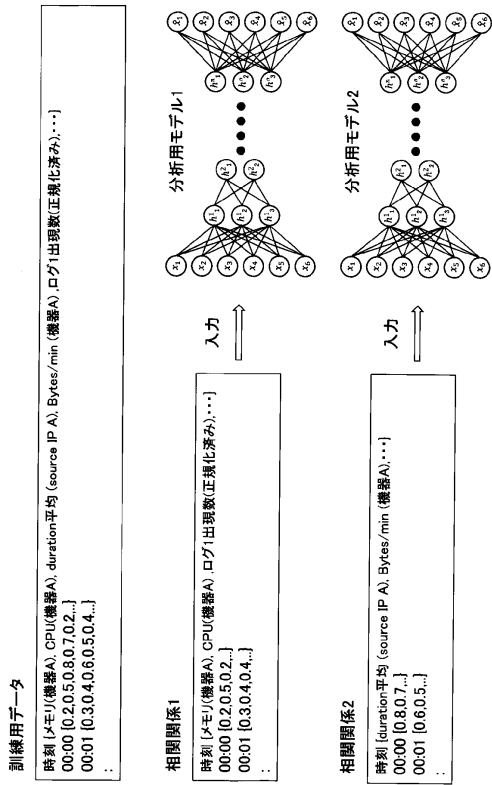
【図3】



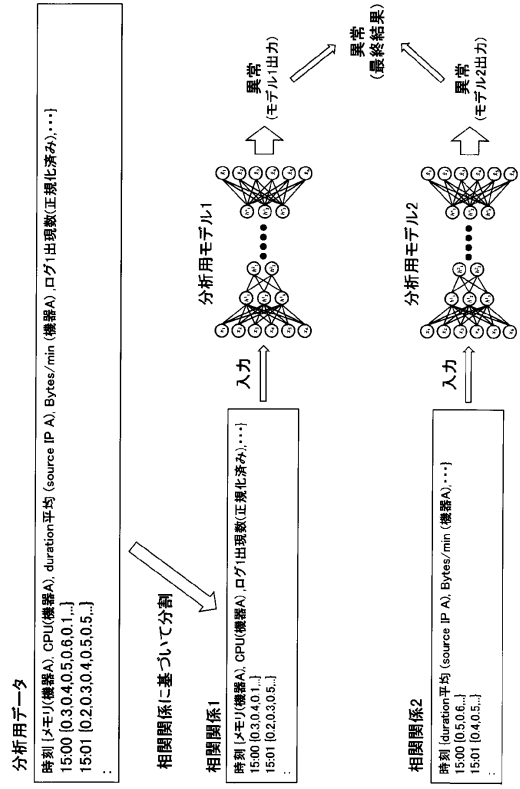
【図4】



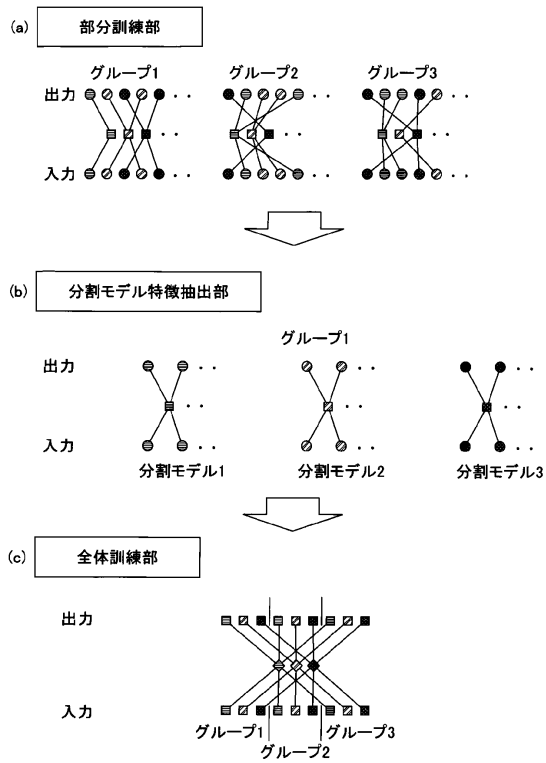
【図 5】



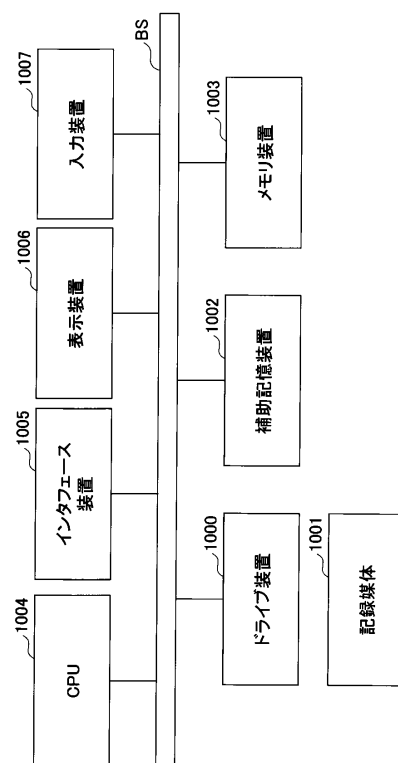
【図 6】



【図 7】



【図 8】



10

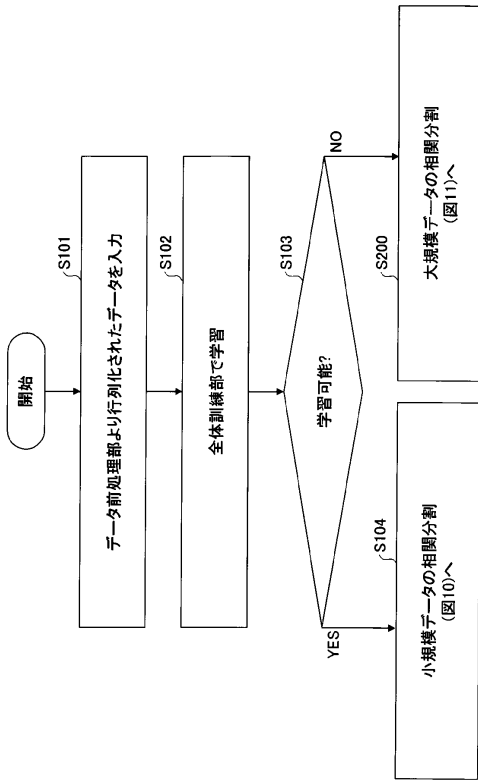
20

30

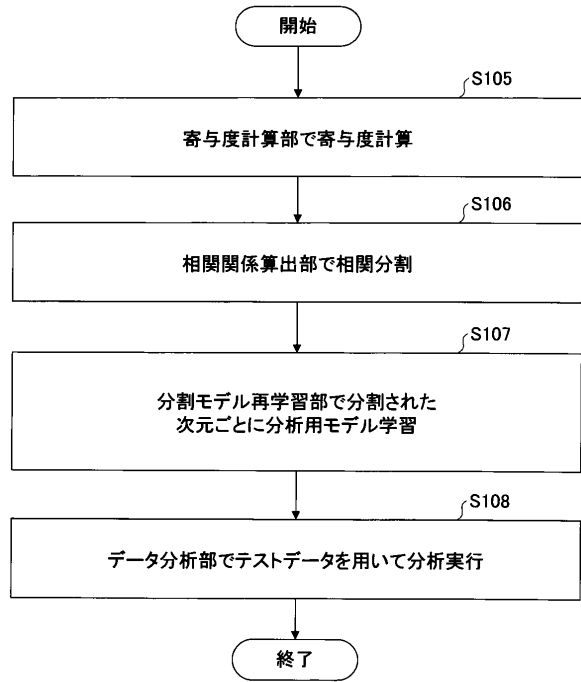
40

50

【図9】



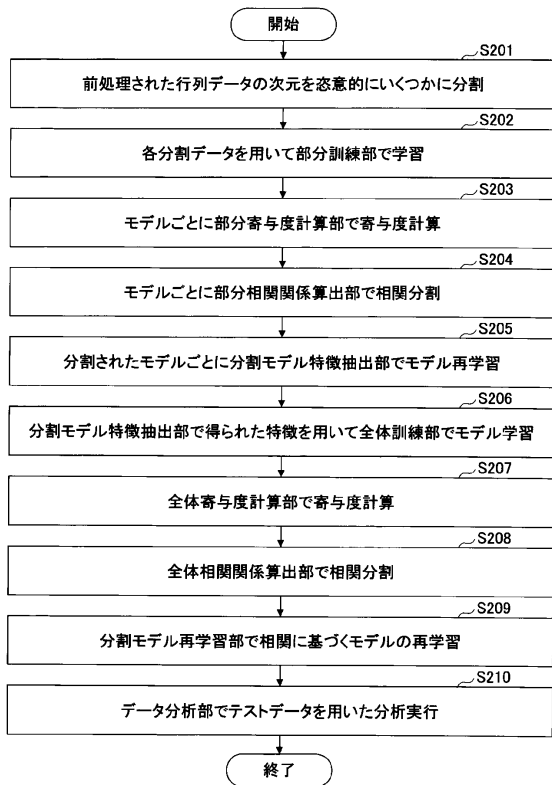
【図10】



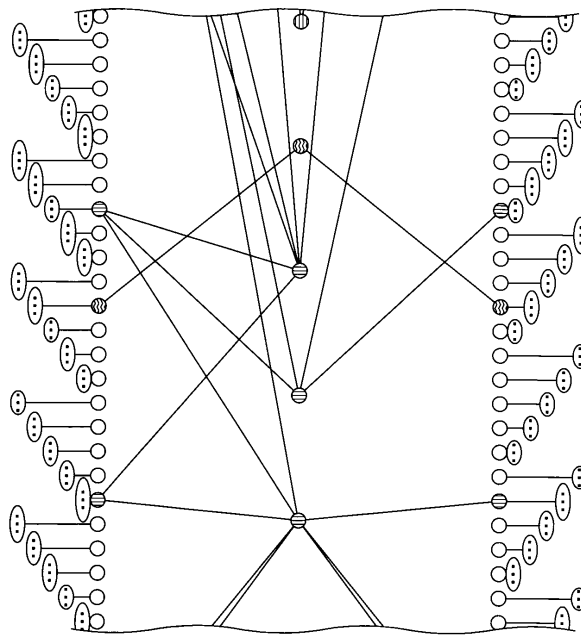
10

20

【図11】



【図12】



30

40

50

【 1 3】

モデル分割関連値 (分割に対する比率)	モデル分割による次元数	AUC
0 (普通のAE)	122 (1モデル)	0.730
0.2	22,1,11,2,1,85 (6モデル)	0.754
0.5	19,1,1,5,5,1,2,1,87 (9モデル)	0.752
1	9,1,1,1,5,1,2,1,1,99 (11モデル)	0.700

10

20

30

40

50

フロントページの続き

審査官 石川 亮

- (56)参考文献 特開 2018 - 169999 (JP, A)
特開 2017 - 123088 (JP, A)
- (58)調査した分野 (Int.Cl., DB名)
- G06N 3/08
G06F 11/07