

(19) 日本国特許庁 (JP)

(12) 特 許 公 報 (B2)

(11) 特許番号

特許第6516660号

(P6516660)

(45) 発行日 令和1年5月22日 (2019.5.22)

(24) 登録日 平成31年4月26日 (2019.4.26)

(51) Int. Cl. F I
GO6N 3/04 (2006.01) GO6N 3/04 190
GO6N 3/08 (2006.01) GO6N 3/08

請求項の数 14 (全 25 頁)

(21) 出願番号	特願2015-230432 (P2015-230432)	(73) 特許権者	000208891
(22) 出願日	平成27年11月26日 (2015.11.26)		KDDI株式会社
(65) 公開番号	特開2017-97693 (P2017-97693A)		東京都新宿区西新宿二丁目3番2号
(43) 公開日	平成29年6月1日 (2017.6.1)	(74) 代理人	100135068
審査請求日	平成30年3月6日 (2018.3.6)		弁理士 早原 茂樹
		(74) 代理人	100141313
			弁理士 辰巳 富彦
		(72) 発明者	呉 剣明
			埼玉県ふじみ野市大原二丁目1番15号
			株式会社KDDI研究所内
		(72) 発明者	住友 亮翼
			埼玉県ふじみ野市大原二丁目1番15号
			株式会社KDDI研究所内

最終頁に続く

(54) 【発明の名称】 異なる周期階層のデータによる学習を行うデータ予測装置、情報端末、プログラム及び方法

(57) 【特許請求の範囲】

【請求項1】

時系列の対象データをもって学習するニューラルネットワーク (NN) を備え、ある時点又は時間範囲での当該対象データを予測する装置であって、

互いに異なる少なくとも2つの時間周期階層の各々に属する時点又は時間間隔に係る対象データ値を特徴量要素として入力する、NNに係る入力層と、少なくとも該入力層から出力される信号を第1の結合荷重で重み付け処理して生成された第1の処理信号を入力する、NNに係る隠れ層とを有し、該隠れ層から出力される信号を第2の結合荷重で重み付け処理して生成された第2の処理信号を出力する入力・隠れ層手段と、

当該第2の処理信号を入力し、少なくとも、予測に係る時点又は時間範囲における当該対象データの値を出力する、NNに係る出力層を有する出力層手段とを有し、

前記入力・隠れ層手段及び前記出力層手段は、少なくとも2つの当該時間周期階層に係る周期性を反映させるべく第1の結合荷重及び第2の結合荷重を更新する再学習を実行可能である

ことを特徴とするデータ予測装置。

【請求項2】

前記入力・隠れ層手段は、前記隠れ層と同一の信号を出力する層としての、NNに係るコンテキスト層を更に有し、

前記隠れ層は、前記コンテキスト層から出力される信号を第3の結合荷重で重み付け処

10

20

理して生成された第3の処理信号を時間遅れの信号として、当該第1の処理信号と共に入力し、

前記入力・隠れ層手段及び前記出力層手段は、少なくとも2つの当該時間周期階層に係る周期性を反映させるべく第3の結合荷重を更新する再学習を実行可能であることを特徴とする請求項1に記載のデータ予測装置。

【請求項3】

前記入力・隠れ層手段及び前記出力層手段は、少なくとも2つの当該時間周期階層に係る周期性を反映させる当該結合荷重の更新を、バックプロパゲーションによって行うことを特徴とする請求項1又は2に記載のデータ予測装置。

【請求項4】

前記入力・隠れ層手段及び前記出力層手段は、当該時系列の対象データのうち学習に未だ使用されていないデータを用いて再学習を実行することを特徴とする請求項1から3のいずれか1項に記載のデータ予測装置。

【請求項5】

第1の時間単位と、所定数の第1の時間単位から構成される第2の時間単位と、この後、順次1つ前の所定数の時間単位から構成される、順次序数を1だけ増分させた時間単位とが設定されており、

少なくとも2つの当該時間周期階層は、第1の時間単位に係る時間を周期とする第1の時間周期階層と、第2の時間単位に係る時間を周期とする第2の時間周期階層と、この後、順次序数を1だけ増分させた時間単位に係る時間を周期とする時間周期階層とのうちの少なくとも2つの時間周期階層を含むことを特徴とする請求項1から4のいずれか1項に記載のデータ予測装置。

【請求項6】

第1の時間単位は「日」であって第1の時間周期階層は「週」に係る周期階層であり、第2の時間単位は「週」であって第2の時間周期階層は「月」に係る周期階層であり、第3の時間単位は「月」であって第3の時間周期階層は「年」に係る周期階層であることを特徴とする請求項5に記載のデータ予測装置。

【請求項7】

当該時間周期階層の各々に属する時点又は時間間隔に係る複数の対象データ値は、第1の時間周期階層を構成する全ての第1の時間単位の中の一部の時間単位に係る対象データ値群と、第2の時間周期階層を構成する全ての第2の時間単位の中の一部の時間単位に係る対象データ値群と、この後、順次序数を1だけ増分させた時間周期階層を構成する全ての時間単位の中の一部の時間単位に係る対象データ値群とのうちの少なくとも2つの群を含むことを特徴とする請求項5又は6に記載のデータ予測装置。

【請求項8】

第1の時間周期階層を構成する全ての第1の時間単位の中の一部の時間単位に係る対象データ値、第2の時間周期階層を構成する全ての第2の時間単位の中の一部の時間単位に係る対象データ値、及び、この後順次序数を1だけ増分させた時間周期階層を構成する全ての時間単位の中の一部の時間単位に係る対象データ値は、自らに係る時間周期階層を構成する時間単位における、第1の時間単位を構成する基本時間単位での値をとることを特徴とする請求項7に記載のデータ予測装置。

【請求項9】

第1の時間単位は「日」であって第1の時間周期階層は「週」に係る周期階層であり、第2の時間単位は「週」であって第2の時間周期階層は「月」に係る周期階層であり、第3の時間単位は「月」であって第3の時間周期階層は「年」に係る周期階層であり、

第1の時間単位、第2の時間単位及び第3の時間単位に係る対象データ値は、当該時間単位における所定の「時刻」又は「時間帯」での値をとることを特徴とする請求項8に記載のデータ予測装置。

【請求項10】

前記出力層手段から出力された予測結果としての当該対象データの値に対する評価結果

10

20

30

40

50

を入力可能なインタフェースを更に備えており、

前記入力・隠れ層手段及び前記出力層手段は、取得された当該評価結果に基づいて生成された教師データを用いて再学習を実行する

ことを特徴とする請求項 1 から 9 のいずれか 1 項に記載のデータ予測装置。

【請求項 1 1】

当該時間周期階層の各々に属する時点又は時間間隔に係る対象データ値は、当該時点又は時間間隔での予測対象における移動予測の基準位置及び移動先の位置に係るデータ値であり、

前記出力層手段によって出力される、予測に係る時点又は時間範囲における当該対象データの値は、当該予測対象における移動先の位置の予測値となる

10

ことを特徴とする請求項 1 から 1 0 のいずれか 1 項に記載のデータ予測装置。

【請求項 1 2】

G P S (Global Positioning System) によって所在位置に係る情報を取得可能な測位部を更に有し、取得された当該所在位置に係る情報に基づいて、前記入力層によって入力される特徴量と、前記再学習に使用される教師データとのうちの少なくとも一方を生成可能な請求項 1 1 に記載のデータ予測装置、であることを特徴とする情報端末。

【請求項 1 3】

時系列の対象データをもって学習するニューラルネットワーク (N N) を備え、ある時点又は時間範囲での当該対象データを予測する装置に搭載されたコンピュータを機能させるプログラムであって、

20

互いに異なる少なくとも 2 つの時間周期階層の各々に属する時点又は時間間隔に係る対象データ値を特徴量要素として入力する、N N に係る入力層と、少なくとも該入力層から出力される信号を第 1 の結合荷重で重み付け処理して生成された第 1 の処理信号を入力する、N N に係る隠れ層とを有し、該隠れ層から出力される信号を第 2 の結合荷重で重み付け処理して生成された第 2 の処理信号を出力する入力・隠れ層手段と、

当該第 2 の処理信号を入力し、少なくとも、予測に係る時点又は時間範囲における当該対象データの値を出力する、N N に係る出力層を有する出力層手段と

してコンピュータを機能させ、

前記入力・隠れ層手段及び前記出力層手段は、少なくとも 2 つの当該時間周期階層に係る周期性を反映させるべく第 1 の結合荷重及び第 2 の結合荷重を更新する再学習を実行可能である

30

ことを特徴とするデータ予測プログラム。

【請求項 1 4】

時系列の対象データをもって学習するニューラルネットワーク (N N) を備え、ある時点又は時間範囲での当該対象データを予測する装置におけるデータ予測方法であって、

互いに異なる少なくとも 2 つの時間周期階層の各々に属する時点又は時間間隔に係る対象データ値を特徴量要素として、N N に係る入力層に入力するステップと、

少なくとも前記入力層から出力される信号を第 1 の結合荷重で重み付け処理して生成された第 1 の処理信号を、N N に係る隠れ層に入力し、該隠れ層から出力される信号を第 2 の結合荷重で重み付け処理して生成された第 2 の処理信号を出力するステップと、

40

当該第 2 の処理信号を、N N に係る出力層に入力し、少なくとも、予測に係る時点又は時間範囲における当該対象データの値を出力するステップと、

少なくとも 2 つの当該時間周期階層に係る周期性を反映させるべく第 1 の結合荷重及び第 2 の結合荷重を更新する再学習を実行するステップとを有することを特徴とするデータ予測方法。

【発明の詳細な説明】

【技術分野】

【0 0 0 1】

本発明は、ユーザの行動履歴に基づき、機械学習を用いて、ユーザに係る行動や状態を予測する情報処理技術に関する。

50

【背景技術】

【0002】

近年、スマートフォンやタブレット型コンピュータ、さらにはカーナビゲーション装置といった高機能の通信端末が普及し、ユーザは、時間や場所を問わず所望の情報を取得することが可能となっている。

【0003】

このような情報へのアクセスの際、ユーザは、その都度、キーワード等の入力やメニュー操作を実行する必要がある、場合によっては相当の手間となる。さらに、例えば、歩行中や運転中にインターネット情報の検索を行ったり、ナビゲーションのルート・目的地の設定を行ったりすることは、操作の手間がかかるだけでなく、危険を伴う場合も少なくない。

10

【0004】

このようなユーザ操作の問題を解決すべく、例えば、音声対話機能を利用して指等による操作を経ずに情報へアクセス可能にしたり、さらには、端末や車に搭載されたGPS (Global Positioning System) 等のセンサを利用してユーザの状況にある程度把握し、ユーザの行動に関する情報をプッシュで提示したりする方式が提案されてきた。しかしながら、このような方式では、利便性が低かったり予測精度の向上が困難であったりといった問題が生じていた。

【0005】

そこで、他の解決策として、ニューラルネットワーク (NN) を利用してユーザに係る行動や状態を予測し、対応する情報を提示する技術が提案されている。例えば、特許文献1は、過去1年間を通した日最大電力需要量実績値のデータに基づき、階層型NNを使用して、日最大電力需要量と気象条件等との非線形な相関関係を学習させて予測モデルを生成し、日最大電力需要量を予測する技術を開示している。ここで、学習には、誤差逆伝播法 (バックプロパゲーション) のアルゴリズムが使用されている。

20

【0006】

また、この特許文献1は、予測精度の向上を意図して、複数年度における実績データに基づき、各年別にNNを使用して学習処理を行い、出力値と電力需要量実績値との相対誤差の中から年増加補正係数を設定し、予測値に対する年増加補正を行う仕組みも提案している。

30

【0007】

さらに、特許文献2は、リカレント型ニューラルネットワーク (RNN) に対し、過去の連続した気象データと当日の気象データとを入力して数時間先の風速を予測し、その情報を基にして風力発電設備の発電電力を予測する技術を提案している。

【先行技術文献】

【特許文献】

【0008】

【特許文献1】特開平9-22402号公報

【特許文献2】特開2007-56686号公報

【発明の概要】

40

【発明が解決しようとする課題】

【0009】

しかしながら、特許文献1に記載されたような階層型NNを用いた従来技術では、学習処理や予測処理を実行するために例えばハイスペックのサーバを必要とし、処理負担が膨大なものとなる。従って、例えば、予測結果に基づく情報を提示すべきユーザの所持する端末や組み込みシステムにおいて、予測情報を生成することは非常に困難となる。

【0010】

この点、特許文献1に記載された技術では、計算量を削減するため、学習用のデータとして1年間の日電力需要量のみを利用し、このような学習による予測値に対し年増加補正係数を用いて予測結果を補正している。しかしながら、過去に遡って全ての実績値データ

50

によって学習した結果と比較すると、予測精度は相当に低下してしまう。

【0011】

一方、特許文献2に記載されたような従来技術では、予測情報の生成にRNNを利用している。RNNは、時系列データを学習し、それに基づく予測値を出力するのに適したシステムである。このRNNは、上述したような階層型NNと比較すると、学習処理の速度は一般により高い。しかしながら、予測精度を向上させるためには、現在から見て時間的に近接した過去の情報を数多く学習する必要がある、また、入力層や中間層の次元数を増加させなければならず、学習に要する時間がかかってしまう。

【0012】

この点、特許文献2は、RNNでの処理時間を短縮するため、前年同時期の既知の気象データのみを学習させることによって、予測精度の向上と予測演算時間の短縮とが可能となると示唆しているが、具体的な基準は何ら示していない。また、こうした季節性を利用した部分的なデータによる学習手法では、季節性を有しない、例えば電力以外のデータの予測に適用することは困難である。

【0013】

そこで、本発明は、相当の予測精度を確保しつつ、従来に比べて学習処理のための演算量をより低減することの可能なデータ予測装置、プログラム及び方法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】

【0014】

本発明によれば、時系列の対象データをもって学習するニューラルネットワーク（NN）を備え、ある時点又は時間範囲での当該対象データを予測する装置であって、

互いに異なる少なくとも2つの時間周期階層の各々に属する時点又は時間間隔に係る対象データ値を特徴量要素として入力する、NNに係る入力層と、少なくともこの入力層から出力される信号を第1の結合荷重で重み付け処理して生成された第1の処理信号を入力する、NNに係る隠れ層とを有し、この隠れ層から出力される信号を第2の結合荷重で重み付け処理して生成された第2の処理信号を出力する入力・隠れ層手段と、

当該第2の処理信号を入力し、少なくとも、予測に係る時点又は時間範囲における当該対象データの値を出力する、NNに係る出力層を有する出力層手段とを有し、

入力・隠れ層手段及び出力層手段は、少なくとも2つの当該時間周期階層に係る周期性を反映させるべく第1の結合荷重及び第2の結合荷重を更新する再学習を実行可能であることを特徴とするデータ予測装置が提供される。

【0015】

この本発明によるデータ予測装置の一実施形態として、

入力・隠れ層手段は、隠れ層と同一の信号を出力する層としての、NNに係るコンテキスト層を更に有し、

隠れ層は、コンテキスト層から出力される信号を第3の結合荷重で重み付け処理して生成された第3の処理信号を時間遅れの信号として、当該第1の処理信号と共に入力し、

入力・隠れ層手段及び出力層手段は、少なくとも2つの当該時間周期階層に係る周期性を反映させるべく第3の結合荷重を更新する再学習を実行可能であることも好ましい。

【0016】

また、本発明によるデータ予測装置における入力・隠れ層手段及び出力層手段は、少なくとも2つの当該時間周期階層に係る周期性を反映させる当該結合荷重の更新を、バックプロパゲーションによって行うことも好ましい。また、入力・隠れ層手段及び前記出力層手段は、当該時系列の対象データのうち学習に未だ使用されていないデータを用いて再学習を実行することも好ましい。

【0017】

さらに、本発明によるデータ予測装置に係る特徴量要素に関して、

第1の時間単位と、所定数の第1の時間単位から構成される第2の時間単位と、この後、順次1つ前の所定数の時間単位から構成される、順次序数を1だけ増分させた時間単位とが設定されており、

少なくとも2つの当該時間周期階層は、第1の時間単位に係る時間を周期とする第1の時間周期階層と、第2の時間単位に係る時間を周期とする第2の時間周期階層と、この後、順次序数を1だけ増分させた時間単位に係る時間を周期とする時間周期階層とのうちの少なくとも2つの時間周期階層を含むことも好ましい。

【0018】

また、上記の特徴を有する特徴量要素に関して、具体的に、第1の時間単位は「日」であって第1の時間周期階層は「週」に係る周期階層であり、第2の時間単位は「週」であって第2の時間周期階層は「月」に係る周期階層であり、第3の時間単位は「月」であって第3の時間周期階層は「年」に係る周期階層であることも好ましい。

10

【0019】

さらに、上記の特徴を有する特徴量要素に関して、当該時間周期階層の各々に属する時点又は時間間隔に係る複数の対象データ値は、第1の時間周期階層を構成する全ての第1の時間単位の中の一部の時間単位に係る対象データ値群と、第2の時間周期階層を構成する全ての第2の時間単位の中の一部の時間単位に係る対象データ値群と、この後、順次序数を1だけ増分させた時間周期階層を構成する全ての時間単位の中の一部の時間単位に係る対象データ値群とのうちの少なくとも2つの群を含むことも好ましい。

20

【0020】

また、上記の特徴を有する特徴量要素に関して、第1の時間周期階層を構成する全ての第1の時間単位の中の一部の時間単位に係る対象データ値、第2の時間周期階層を構成する全ての第2の時間単位の中の一部の時間単位に係る対象データ値、及び、この後順次序数を1だけ増分させた時間周期階層を構成する全ての時間単位の中の一部の時間単位に係る対象データ値は、自らに係る時間周期階層を構成する時間単位における、第1の時間単位を構成する基本時間単位での値をとることも好ましい。

【0021】

さらに、上記の特徴を有する特徴量要素に関して、具体的に、

第1の時間単位は「日」であって第1の時間周期階層は「週」に係る周期階層であり、第2の時間単位は「週」であって第2の時間周期階層は「月」に係る周期階層であり、第3の時間単位は「月」であって第3の時間周期階層は「年」に係る周期階層であり、

30

第1の時間単位、第2の時間単位及び第3の時間単位に係る対象データ値は、当該時間単位における所定の「時刻」又は「時間帯」での値をとることも好ましい。

【0022】

さらに、本発明によるデータ予測装置における他の実施形態として、

出力層手段から出力された予測結果としての当該対象データの値に対する評価結果を入力可能なインタフェースを更に備えており、

入力・隠れ層手段及び出力層手段は、取得された当該評価結果に基づいて生成された教師データを用いて再学習を実行することも好ましい。

40

【0023】

また、本発明によるデータ予測装置における対象データに関して、

当該時間周期階層の各々に属する時点又は時間間隔に係る対象データ値は、当該時点又は時間間隔での予測対象における移動予測の基準位置及び移動先の位置に係るデータ値であり、

出力層手段によって出力される、予測に係る時点又は時間範囲における当該対象データの値は、当該予測対象における移動先の位置の予測値となることも好ましい。

【0024】

50

また、このような対象データを取り扱うデータ装置が、GPS (Global Positioning System) によって所在位置に係る情報を取得可能な測位部を更に有し、取得された当該所在位置に係る情報に基づいて、入力層によって入力される特徴量と、再学習に使用される教師データとのうちの少なくとも一方を生成可能である情報端末であることも好ましい。

【0025】

本発明によれば、また、時系列の対象データをもって学習するNNを備え、ある時点又は時間範囲での当該対象データを予測する装置に搭載されたコンピュータを機能させるプログラムであって、

互いに異なる少なくとも2つの時間周期階層の各々に属する時点又は時間間隔に係る対象データ値を特徴量要素として入力する、NNに係る入力層と、少なくともこの入力層から出力される信号を第1の結合荷重で重み付け処理して生成された第1の処理信号を入力する、NNに係る隠れ層とを有し、この隠れ層から出力される信号を第2の結合荷重で重み付け処理して生成された第2の処理信号を出力する入力・隠れ層手段と、

当該第2の処理信号を入力し、少なくとも、予測に係る時点又は時間範囲における当該対象データの値を出力する、NNに係る出力層を有する出力層手段としてコンピュータを機能させ、

入力・隠れ層手段及び出力層手段は、少なくとも2つの当該時間周期階層に係る周期性を反映させるべく第1の結合荷重及び第2の結合荷重を更新する再学習を実行可能であることを特徴とするデータ予測プログラムが提供される。

【0026】

本発明によれば、さらに、時系列の対象データをもって学習するNNを備え、ある時点又は時間範囲での当該対象データを予測する装置におけるデータ予測方法であって、

互いに異なる少なくとも2つの時間周期階層の各々に属する時点又は時間間隔に係る対象データ値を特徴量要素として、NNに係る入力層に入力するステップと、

少なくとも入力層から出力される信号を第1の結合荷重で重み付け処理して生成された第1の処理信号を、NNに係る隠れ層に入力し、この隠れ層から出力される信号を第2の結合荷重で重み付け処理して生成された第2の処理信号を出力するステップと、

当該第2の処理信号を、NNに係る出力層に入力し、少なくとも、予測に係る時点又は時間範囲における当該対象データの値を出力するステップと、

少なくとも2つの当該時間周期階層に係る周期性を反映させるべく第1の結合荷重及び第2の結合荷重を更新する再学習を実行するステップとを有するデータ予測方法が提供される。

【発明の効果】

【0027】

本発明のデータ予測装置、情報端末、プログラム及び方法によれば、相当の予測精度を確保しつつ、従来に比べて学習処理のための演算量をより低減することができる。

【図面の簡単な説明】

【0028】

【図1】本発明によるデータ予測装置の一使用形態を示す模式図である。

【図2】本発明によるデータ予測装置の一実施形態における機能構成を示す機能ブロック図である。

【図3】本発明によるデータ予測装置の他の実施形態における機能構成を示す機能ブロック図である。

【図4】深層学習部における学習・データ予測処理の一実施形態を示す模式図である。

【図5】本発明に係る特徴量ベクトルの生成の一実施形態を示す模式図である。

【図6】図4に示された各ステップの一実施形態を実行するためのプログラムコードを示す模式図である。

【図7】図4に示された各ステップの一実施形態を実行するためのプログラムコードを示す模式図である。

【図8】本発明によるデータ予測方法の一実施例及び比較例を示すテーブルである。

【発明を実施するための形態】

【0029】

以下、本発明の実施形態について、図面を用いて詳細に説明する。

【0030】

図1は、本発明によるデータ予測装置の一使用形態を示す模式図である。

【0031】

図1によれば、本発明のデータ予測装置であるスマートフォン1は、移動体としての自動車内に設置されたクレードル8に取り付けられ、カーナビゲーション機能を含む情報提供用の端末として使用されている。図1では、スマートフォン1は、運転席と助手席との真ん中前方の位置であっての運転席からよく見える位置に配置されている。このように、スマートフォン1は、自動車と電氣的に接続させる必要がなく簡便な形で設置される。

10

【0032】

ここで、本発明によるデータ予測装置は、当然にスマートフォンに限定されるものではない。例えば、タブレット型コンピュータやウェアラブル端末といった移動携帯端末であってもよい。さらには、カーナビゲーションユニットやパーソナル・コンピュータ（PC）等の自動車に積み込み可能な情報機器であってもよい。いずれにしても、このような機器は、例えば、以下に説明する本発明の主要機能を作動させるアプリケーション・プログラムを機器内コンピュータに実行させることによって、本発明のデータ予測装置としての機能を発揮するものとすることができる。

【0033】

20

スマートフォン1は、時系列の対象データをもって学習するニューラルネットワーク（NN）を備えており、ある時点又は時間範囲での対象データの値を予測する。本実施形態では、スマートフォン1は、GPS（Global Positioning System）によって所在位置に係る情報を取得可能な測位部102を有し、測位部102で取得された所在位置に係る情報に基づいて、所定時間経過後の移動先（予測すべき時点又は時間範囲での所在位置）を予測することができる。また、生成した移動先の予測結果に基づき、この結果に関連するコンテンツを取得又は選択し、ユーザに提供可能であることも好ましい。

【0034】

ここで、スマートフォン1における特徴的な具体的構成を挙げると、最初に、備えられたNNへ入力する特徴量ベクトルとして、

30

（A）互いに異なる「少なくとも2つの時間周期階層」の各々に属する時点又は時間間隔に係る対象データ値を特徴量要素とする特徴量ベクトルを採用している。

【0035】

この特徴量ベクトルについては、後に図5を用いて詳細に説明するが、例えば、2つの異なる「時間周期階層」として「1週間」及び「1ヶ月」を採用し、

（ア）「1週間」に属する3つの時点（時間間隔）として「1日前」、「2日前」及び「3日前」を採用し、

（イ）「1ヶ月」に属する3つの時点（時間間隔）として「1週間前」、「2週間前」及び「3週間前」を採用する。

40

ここで、各時点でのデータ値を要素とした特徴量ベクトルを生成することができる。

【0036】

また、スマートフォン1は、

（B）上記（A）の特徴量ベクトルを入力する、NNに係る「入力層」と、少なくともこの「入力層」から出力される信号を第1の結合荷重で重み付け処理して生成された第1の処理信号を入力する、NNに係る「隠れ層」とを有し、この「隠れ層」から出力される信号を第2の結合荷重で重み付け処理して生成された第2の処理信号を出力する入力・隠れ層部と、

（C）第2の処理信号を入力し、少なくとも、予測に係る時点又は時間範囲における対象データの値を出力する、NNに係る「出力層」を有する出力層部と

50

を有しており、

(D) これらの入力・隠れ層部及び出力層部は、「少なくとも2つの時間周期階層」に係る周期性を反映させるべく第1の結合荷重及び第2の結合荷重を更新する再学習を実行可能である

ことを特徴としている。

【0037】

上記構成(C)にあるように、「出力層」は、予測に係る時点又は時間範囲における対象データの値を出力する。例えば、測位部102で取得されたスマートフォン1の所在位置に係る情報に基づいて特徴量ベクトルを生成し、「入力層」に入力した場合、「出力層」は、予測に係る時点又は時間範囲での所在位置(移動先)の予測値を出力することができる。

10

【0038】

また、例えば、測位部102で取得されたスマートフォン1の所在位置に係る情報に基づいて教師データを生成し、上記構成(D)における再学習を実行してもよい。この場合、スマートフォン1のユーザの移動先に見られる「少なくとも2つの時間周期階層」に係る周期性、例えば毎週所定の曜日に所定場所に出向いたり、又は毎月下旬に所定の場所でミーティングを行ったりといった移動行動に見られる周期性を、第1の結合荷重及び第2の結合荷重に反映させることができるのである。これにより、学習処理のための演算量を抑えつつ、十分な予測精度を確保することが可能となる。

【0039】

20

さらに、スマートフォン1においては、

(E) 入力・隠れ層部は、「隠れ層」と同一の信号を出力する層としての、NNに係る「コンテキスト層」を更に有し、

(F) 「隠れ層」は、「コンテキスト層」から出力される信号を第3の結合荷重で重み付け処理して生成された第3の処理信号を時間遅れの信号として、第1の処理信号と共に入力し、

(G) 入力・隠れ層部及び出力層部は、「少なくとも2つの時間周期階層」に係る周期性を反映させるべく第3の結合荷重を更新する再学習を実行可能である

ことも好ましい。

【0040】

30

この場合、この機械学習機能部は、リカレント型ニューラルネットワーク(RNN)に相当する。特に、一次遅れの「コンテキスト層」からの出力に対しても結合荷重を持たせているので、この機能部は、データの有する複数の「時間周期階層」を適切に反映可能な内部記憶部分を有することになる。このように、「隠れ層」の出力の一部が入力層側にフィードバックしており、しかも結合状態が非対称である機能部は、入力となる時系列データの有する周期性及びその変化を的確に識別することに適している。

【0041】

ここで、本発明のスマートフォン1は、特に、上記構成(A)の、互いに異なる「少なくとも2つの時間周期階層」の各々に係る対象データ値を特徴量要素とする特徴量ベクトルを採用している。その結果、対象データが時系列で連続しておらず離散的であってより少ないデータ量であったとしても、各「時間周期階層」でのデータ値を取り込むことによって、対象データの有する様々な周期性及びその変化を、効率良く学習することができる。これにより、相当の予測精度を確保しつつも、従来に比べて学習処理のための演算量をより低減することが可能となるのである。

40

【0042】

さらに、上記構成(E)～(G)を有するRNNを用いて学習を行う場合、例えば時系列データを取り扱う際に従来よく用いられてきた隠れマルコフモデル(HMM)と比較しても、このHMMのように先験的知識を必要としない。即ち、このような先験的知識を用いることなく、例えばユーザの移動履歴データのみに基づき、過去の時系列データと現在の時系列データとを総合した情報解析を可能なフィードバック結合を利用した深層学習を

50

実行することができるのである。これにより、NNでは通常トレードオフ関係にある、高い予測精度と少ない演算量（短い計算処理時間）とを両立することも可能となる。また、従って、サーバ等のハイスpek機器ではなく、スマートフォンや車載カーナビ装置等の情報端末を、本発明のデータ予測装置とすることもできるのである。

【0043】

特に、従来、RNN等の深層学習（Deep Learning）はインターネット上のサーバで実行されるケースが大部分であったが、上述したように各「時間周期階層」でのデータ値に基づいて特徴量を生成することによって演算の収束を早め、学習に必要な計算量を抑制することができるので、近年高スペック化の進んでいる移動情報端末において実行することが可能となるのである。

10

【0044】

ちなみに、上記のHMMでは、状態遷移トポロジの構成設計に先験的知識が必要となるのであり、設計段階ではすべての状態遷移を網羅的にカバーすることができない（例えば、設計段階でトポロジに含まれていない状態遷移の振舞いが認識できない）との問題が生じてしまう。従って、例えば、ユーザの移動先を予測する際、何種類の移動パターンがどのように時系列中に含まれているのかを事前に把握できないため、各隠れ状態間の遷移トポロジを事前に定めることが非常に困難となってしまう。さらに、HMMでは、新しい隠れ状態の追加につれ、それに応じて連続時間のモデルの再学習を最初から行う必要が生じ、これにより計算量が膨大となってしまうので、例えば、ユーザの所持する情報携帯端末を用いてオンラインで適応学習処理を行うといったシステムを構成することも極めて難しいのである。

20

【0045】

ここで、スマートフォン1の具体的使用例を示すと、例えば、過去の移動履歴から予測された所定時間後の移動先の予測結果に基づいて、予測される移動先を示すだけでなく、移動先又はそこに到着するまでの経路上にあるお出かけスポット情報を車内のユーザに提供してもよい。また、移動先近辺の駐車場等、種々の案内情報を車内のユーザに提供することもできる。さらに、予測される移動先が通勤先の会社である場合、会社に係る情報、例えば、当日の仕事のスケジュールを提示してもよい。また、移動先が店舗の場合、この店舗でのタイムサービス情報等を提供することも可能となる。

【0046】

30

このように、予測結果としての移動先に適合した、予め設定されたコンテンツ等の情報を、ユーザ（運転者及び／又は同乗者）に提示することができる。ここで、この提示情報は、個々のユーザに特有のパーソナライズされた情報としてもよい。尚、上記のような情報提示の際には、後に詳細に説明する、GPS（Global Positioning System）による現在位置情報や、VICS（登録商標）（Vehicle Information and Communication System）による渋滞や交通規制等の道路交通情報を取得して利用することも可能である。

【0047】

尚、スマートフォン1等の本発明によるデータ予測装置が設置される移動体は、当然に自動車に限定されない。例えば自転車、歩行者、自動車以外の乗用車、トラック、バス、エレベータ等、所定時間（期間）後の移動先が必ずしも決まっていらず予測することに意義のある移動体であれば、種々のものが対象となり得る。

40

【0048】

さらに、本発明によるデータ予測装置によって予測される予測対象データは、当然に、移動先（所在位置）に限定されるものではない。例えば、株価や為替レート等の金融データ、電力やガス、水道等の使用量データ、さらにはイベントの開催や事件の発生等の社会現象に係るデータ等、時系列のデータであって何らかの周期性の見られる可能性を有するデータであれば、本発明に係る予測対象データとなり得る。

【0049】

[データ予測装置の機能構成]

図2は、本発明によるデータ予測装置の一実施形態における機能構成を示す機能ブロッ

50

ク図である。

【0050】

図2によれば、データ予測装置であるスマートフォン1は、通信インタフェース部101と、測位部102と、ユーザインタフェースとしてのタッチパネル・ディスプレイ（TP/DP）103、スピーカ104及びマイク105と、ログ記憶部106と、結合荷重記憶部107と、コンテンツ蓄積部108と、プロセッサ・メモリとを有する。ここで、プロセッサ・メモリは、スマートフォン1に搭載されたコンピュータを機能させるプログラムを実行することによって、データ予測機能を実現させる。

【0051】

また、プロセッサ・メモリは、位置情報取得部111と、状況取得部112と、特徴量生成部113と、深層学習部12と、予測結果処理部114と、ユーザインタフェース（IF）制御部115とを有する。ここで、ユーザIF制御部115は、対話処理部115cを有することも好ましい。尚、図2によれば、各機能構成部を矢印で接続した処理の流れは、本発明によるデータ予測方法の一実施形態としても理解される。

【0052】

図2において、通信インタフェース部101は、

（a）民間で又は公的に提供された気象情報サーバ6から、現在又は所定時間（期間）後の気象情報を受信したり、

（b）VICS（登録商標）サーバ7から、交通又は運行情報を配信する交通又は運行情報（例えばVICS（登録商標）情報）を受信したり

することができる。また、

（d）予測結果処理部114において決定された、移動先の予測結果又はこの結果に応じて決定された提示情報を、外部の装置、例えば車内に設置されたディスプレイ装置やスピーカ装置に送信する

ことも好ましい。

【0053】

測位部102は、GPSシステムをなすGPS衛星5からの測位電波を捕捉し、スマートフォン1（自動車）の現在位置を測定する。また、位置情報取得部111は、測位部102での測定結果を受けて現在位置情報、例えば現在位置の緯度及び経度の情報を取得する。

【0054】

状況取得部112は、通信インタフェース部101で受信された気象情報や交通又は運行情報を取得し、現在位置（現在時点）における気象状態や交通又は運行状態を決定し、特徴量生成部113での特徴量生成のための情報を提供する。

【0055】

特徴量生成部113は、互いに異なる「少なくとも2つの時間周期階層」の各々に属する時点又は時間間隔に係る対象データ値を特徴量要素とする特徴量ベクトルを生成し、深層学習部12に出力する。特徴量ベクトルの具体的な構成については、後に図5を用いて詳細に説明する。

【0056】

深層学習部12は、

（a）入力層121i、隠れ層121h及びコンテキスト層（CON層）121cを有する入力・隠れ層部121と、

（b）出力層122oを有する出力層部122とを有する。

【0057】

ここで、入力層121iは、互いに異なる少なくとも2つの時間周期階層の各々に属する時点又は時間間隔に係る対象データ値を特徴量要素として入力する層である。また、隠れ層121hは、

（a1）入力層121iから出力される信号を第1の結合荷重で重み付け処理して生成さ

10

20

30

40

50

れた第1の処理信号

を入力し、さらに、

(a2) コンテキスト層121cから出力される信号を第3の結合荷重で重み付け処理して生成された第3の処理信号

を時間遅れの信号として上記の第1の処理信号と共に入力する。

【0058】

また、入力・隠れ層部121は、隠れ層121hから出力される信号を第2の結合荷重で重み付け処理して生成された第2の処理信号を、出力層部122に出力する。出力層部122の出力層122oは、第2の処理信号を入力し、少なくとも、予測に係る時点又は時間範囲における対象データの値を出力する。

10

【0059】

さらに、入力・隠れ層部121及び出力層部122は、少なくとも2つの時間周期階層に係る周期性を反映させるべく第1の結合荷重、第2の結合荷重及び第3の結合荷重を更新する再学習を、誤差逆伝播法（バックプロパゲーション）によって実行可能なように構成されている。

【0060】

ここで、時間周期階層の各々に属する時点又は時間間隔に係る対象データ値が、これらの時点又は時間間隔での予測対象における移動予測の基準位置及び移動先の位置に係るデータ値である場合、出力層部122によって出力される、予測に係る時点又は時間範囲における当該対象データの値は、予測対象における移動先の位置の予測値となる。尚、深層学習部12における学習・予測処理については、後に図4を用いて詳細に説明する。

20

【0061】

予測結果処理部114は、出力層部122から出力された、予測に係る時点又は時間範囲における対象データの値を予測結果に加工し、この予測結果の情報をユーザIF制御部115に出力する。例えば、出力層122o（出力層部122）からの出力予測値が0から1までの値で規格化された数値である場合、この数値を（移動先等の）予測結果に変換する。ここで、予測結果処理部114は、予測数を設定し、深層学習部12に対し、予測処理をこの予測数だけ繰り返し行わせ、所定時間（期間）を1単位として、現在から1単位分だけ先での予測値だけでなく、複数単位分先での予測値を算出させて、これに応じた予測結果を生成することも好ましい。

30

【0062】

さらに、予測結果処理部114は、コンテンツ蓄積部108から、生成した予測結果に対応したコンテンツを取り出し、ユーザIF制御部115を介して外部に提示させてもよい。例えば、TP/DP103に、コンテンツに係る提示情報を表示させてもよく、スピーカ104に音声としての提示情報を出力させてもよい。ここで、コンテンツ記憶部108は、提示情報（コンテンツ）を外部のコンテンツサーバから通信インタフェース101を介して取得してもよい。また、ユーザによるTP/DP103を介した入力操作によって、所定の提示情報を入力してもよい。

【0063】

ユーザIF制御部115は、予測結果処理部114から入力した予測結果情報やコンテンツを提示情報としてTP/DP103に表示させたり、対話処理部115cを介してスピーカ104から提示情報に対応する音声出力させたりする。また、ユーザIF制御部115は、提示された予測結果に対するユーザによる肯定や否定といった評価結果である応答情報を、TP/DP103に対する操作を介して、マイク105への音声入力及び対話処理部115cを介して、又はリモコン等の外部機器を介して取得し、予測結果処理部114に出力する。

40

【0064】

この際、予測結果処理部114は、このユーザによる評価結果（応答情報）に基づき、再学習のための教師データを生成し、深層学習部12に出力して再学習を実行させる。即ち、予測結果処理部114は、NNシステムの信頼性を高めるため、実環境の利用におい

50

てユーザの肯定・否定の応答を用いて学習データを更新する手段としても機能する。尚、予測値が出力され予測結果が提示された後に、ユーザから何ら応答が入力されない場合、現状で使用された特徴量および予測値をログ記憶部106に記憶する。一方、例えばユーザから否定（異議）が入力された場合、予測値をユーザからの応答に基づいて生成された修正情報に置き換えて、使用された特徴ベクトルとともにログ記憶部106に書き込む。

【0065】

ログ記憶部106は、特徴量ベクトルを生成するためのユーザの移動履歴データ（例えば、日時とその時点（期間）での所在位置とを対応付けた時系列データ）を記憶する。また、ユーザによる評価結果（応答情報）に基づいて、学習すべき移動履歴データを更新する機能も備えている。

10

【0066】

結合荷重記憶部107は、深層学習部12において学習処理の際に算出・更新された結合荷重（第1、第2及び第3の結合荷重）と、ログ記憶部106に記憶された学習済みである最後のレコードのレコードIDとを記憶する。また、再学習の際、深層学習部12からの読み出し指示に応じ、記憶した結合荷重及びレコードIDを深層学習部12に出力する。ここで、この再学習は、読み出されたレコードIDの次のIDを有する新規レコードから行われることも好ましい。この場合、言い換えると、学習に未だ使用されていないレコードを用いて再学習を実行することになる。このように履歴の差分のみに対して再学習を行うことによって、学習処理に必要なとなる演算量をより低減することが可能となるのである。

20

【0067】

図3は、本発明によるデータ予測装置の他の実施形態における機能構成を示す機能ブロック図である。

【0068】

図3によれば、本発明のデータ予測装置としてのデータ予測サーバ2と、スマートフォン3とが、例えばアクセスネットワーク及びインターネットを介して通信接続されることによってデータ予測システムが構築されている。

【0069】

このうち、データ予測サーバ2は、それぞれスマートフォン1（図2）における対応する機能部相当の手段である、通信インタフェース部201及び通信制御部211と、ログ記憶部206と、結合荷重記憶部207と、深層学習部22とを有している。

30

【0070】

また、スマートフォン3は、同じくそれぞれスマートフォン1（図2）における対応する機能部相当の手段である、通信インタフェース部301及び通信制御部312と、測位部302及び位置情報取得部311と、ユーザインタフェースとしてのユーザIF部303及びユーザIF制御部315と、特徴量生成部313と、予測結果処理部314とを有している。

【0071】

即ち、本データ予測システムは、図2に示したスマートフォン1における深層学習機能を外部のサーバに持たせ、このサーバをデータ予測装置とした実施形態になっている。ここで、スマートフォン3で生成された特徴量ベクトルが、データ予測サーバ2に送信される。一方、データ予測サーバ2で生成された予測値が、スマートフォン3に送信され、スマートフォン3の予測結果処理部314において予測結果に変換されて、ユーザIFを介してユーザにこの予測結果が提示される。さらに、ユーザIFを介して入力されたユーザによる肯定又は否定の評価結果に基づいて、スマートフォン3の予測結果処理部314が教師データを生成し、この教師データがデータ予測サーバ2で再学習に使用されてもよい。

40

【0072】

ここで、特徴量生成部313及び／又は予測結果処理部314を、スマートフォン3ではなく、データ予測サーバ2が備えている形態も可能である。いずれにしても、以上に述

50

べたような構成のデータ予測システムにおいて、互いに異なる「少なくとも2つの時間周期階層」の各々に属する時点又は時間間隔に係る対象データ値を要素とする特徴量ベクトルを採用することによって、相当の予測精度を確保しつつ、従来に比べて学習処理のための演算量をより低減することが可能となる。

【0073】

また、例えば、レンタカーサービス事業に本データ予測システムを適用し、レンタカーのユーザが、自ら所有するスマートフォンを、日頃使い慣れたカーナビ装置としてこのレンタカーで利用することが可能となる。この際、コールセンターに設置されたデータ予測サーバ2に、ユーザ毎の移動履歴及び教師データによって学習を行う、ユーザ毎に構築された深層学習部22を用意し、算出された予測移動先（予測値）を、該当するユーザの使用しているレンタカーにダウンロードさせて、この予測結果を利用させることも可能となる。

10

【0074】

さらに、データ予測サーバ2相当のデータ予測装置が、自動車（移動体）の内部に組み込まれている実施形態も可能である。例えば、自動車に搭載された制御用CPUにおいて、他の標準的な機能と並行して本発明のデータ予測機能を実現する場合、このCPUを含む自動車内のユニットが、本発明のデータ予測装置に該当することになる。この場合、この車内ユニットと、ユーザが車内に持ち込んだスマートフォン等とが、有線又は無線で通信接続されることも好ましい。

【0075】

20

図4は、深層学習部12における学習・データ予測処理の一実施形態を示す模式図である。

【0076】

図4に示すように、深層学習部12は、入力層121i、隠れ層121h、コンテキスト層121c及び出力層121oによってRNNを構成している。RNNは、時系列データを取り扱うのに非常に適したネットワークである。以下、実施される深層学習の機能を分かり易く説明することを目的として、深層学習部12が、RNNの一種である単層構造のエルマンネットワーク（Elman Network）であるとして説明を行う。

【0077】

本来、ニューラルネットワーク（NN）は、人間の脳神経回路の仕組みを模したモデルであって、入力層、隠れ層（中間層）及び出力層をモデル構成要素とする。これらの各層は、脳内の神経細胞に相当する複数のニューロン（Neuron）を備えている。各ニューロンの基本的な機能は、信号の入力及び出力である。但し、1つの層のニューロンから次の層のニューロンへ信号伝達を行う際、1つの層のニューロンから出力された信号をそのまま、次の層のニューロンが入力するのではなく、出力された信号に結合荷重をもって重み付けし、重み付けされた出力信号の総和が各ニューロンについて設定されている閾値を超えた場合にのみ、次の層のニューロンへ信号を送出する。ここで、教師データを用いた学習からこれらニューロン間の結合荷重を算出・更新し適用していくことによって、予測対象となる新たなデータを入力した際に、予測値の出力が可能となるのである。

30

【0078】

図4に示したエルマンネットワーク（深層学習部12）は、ループ構造を含む深層NNであり、入力層121i、隠れ層121h、コンテキスト層121c及び出力層121oを備えている。入力層121iは、入力信号である特徴量ベクトルを処理するn個の入力層ニューロン n_i を有する。この入力層ニューロン n_i の数nは、特徴量ベクトルの次数（要素数）と一致する。コンテキスト層121cは、直前の隠れ層121hの状態を複製したm個のコンテキスト層ニューロン n_c を備えている。このコンテキスト層ニューロン n_c の数mは、次に述べる隠れ層ニューロン n_h の数と一致する。

40

【0079】

隠れ層121hは、入力層ニューロン n_i とコンテキスト層ニューロン n_c からの入力信号を処理するm個の隠れ層ニューロン n_h を備えている。この隠れ層ニューロン n_h の

50

数 m は任意でありネットワーク設計パラメータの1つとなる。出力層121oは、隠れ層ニューロン n_h からの入力信号を処理する出力層ニューロン n_o を備えている。この出力層ニューロン n_o の数 k は、予測値を構成する予測ベクトルの次元数と一致する。例えば、予測値が移動先（予測所在位置）の位置情報のみであれば、 $k=1$ とすることができる。

【0080】

上述したニューロンはそれぞれの活性化関数を有し、ニューロン毎に異なるパラメータを保持している。ニューロンは、例えばバックプロパゲーションによる学習過程を実行することによって、保持するパラメータを調整し、活性化関数の表現を決定して、上記の結合荷重の算出・更新を行う。この学習過程は何度も繰り返される。即ち、データを入力した結果出力された予測データ（学習値）と教師データ（実測値）とから算出される誤差を、後ろ向きに伝搬して上記のパラメータを微調整し、さらに次のデータを入力した結果出力されたデータによる誤差を、再度後ろ向きに伝搬して再度パラメータを微調整する、といったように誤差が極小（最小）となり収束するまで、このパラメータの微調整処理を実行し続ける。次いで、予測処理においては、例えば、測定される現在位置やその日時といった即時状況の値をNN（深層学習部12）に入力することによって、予測値、例えば移動先（予測される所在位置）が出力されるのである。

【0081】

図4に示したRNN（エルマンネットワーク）は、以上に述べた構成を有するが故に、比較的生体（脳）に近い特長を有し、過去との相関を強く有する人間（脳）の行動パターンを予測するのに非常に適している。具体的には、隠れ層ニューロン n_h が、コンテキスト層ニューロン n_c を介し、隣り合う段階の隠れ層ニューロン n_h と相互に影響を及ぼしながら状態を更新することによって、予測対象の行動パターンが、より網羅的且つ的確に結合荷重に反映されていくのである。このように、図4の深層学習部12はリカレント型であるので、隠れ層及びコンテキスト層が展開・循環して機能することによって多層的な深層学習を行うことが可能となる。また、この展開・循環処理を繰り返すことによって、予め期待された時間周期以外の周期パターンを認識することも可能となるのである。

【0082】

尚、上述した隠れ層ニューロン n_h の数（隠れ層121hのユニット数） m 、学習過程の繰り返し回数、及びパラメータ微調整における調整幅（学習率）は、ネットワーク設計事項であり、経験値として決定してもよい。

【0083】

以後、図4に示した、データ予測方法におけるステップ1～ステップ7の具体的内容を、図5～7を用いて詳細に説明する。

【0084】

[データ予測方法]

図5は、本発明に係る特徴量ベクトルの生成の一実施形態を示す模式図である。

【0085】

図5（A）には、ステップ1（図4）で入力層121iに入力される特徴量ベクトルの一実施形態が示されている。この特徴量ベクトルは、ユーザの過去の移動履歴の時系列データ（例えば、日時と当該日時での所在位置（出発地）と目的地とを対応付けた時系列データ）を用いて、ユーザが何時どこからどこまで移動したのかを分析することによって生成される。

【0086】

本実施形態の特徴量ベクトルは、具体的に、

1. 対象移動履歴の出発時の日付
2. 対象移動履歴の出発時の曜日
3. 対象移動履歴の出発時の平日・祝日区分
4. 対象移動履歴の出発時の時間帯区分
5. 対象移動履歴の出発時の位置区分

6. 複数の時間周期階層に基づいた時系列の移動先

- 6 a・当日（直前）の移動先
- 6 b・1 日前の同時帯（上記 4. の時間帯）の移動先
- 6 c・2 日前の同時帯の移動先
- 6 d・3 日前の同時帯の移動先
- 6 e・1 週間前の同時帯の移動先
- 6 f・2 週間前の同時帯の移動先
- 6 g・3 週間前の同時帯の移動先
- 6 h・1 ヶ月前の同時帯の移動先
- 6 i・2 ヶ月前の同時帯の移動先
- 6 j・3 ヶ月前の同時帯の移動先
- 6 k・6 ヶ月前の同時帯の移動先
- 6 l・1 年前の同時帯の移動先

10

とのベクトル要素を有する。これらのベクトル要素の数は、入力層ニューロン n_i の数 n と一致している。具体的に、ベクトル要素数 n は、上記項目 1.～5. までの 5 つと、上記項目 6. における p ($=12$) 個との合計であって、

$$n = 5 + p = 17$$

となっている。ちなみに、ベクトル要素の値は、0 から 1 までの値に正規化されていることも好ましい。

【0087】

20

(a) ここで、特徴量ベクトルの要素 6 a～6 d においては、第 1 の時間単位を「日」とし、第 1 の時間周期階層は、所定数（7 つ）の「日」から構成される「週」に係る周期階層となっている。

(b) また、特徴量ベクトルの要素 6 e～6 g においては、第 2 の時間単位を「週」とし、第 2 の時間周期階層は、（分数を含む）所定数の「週」から構成される「月」に係る周期階層となっている。

(c) さらに、特徴量ベクトルの要素 6 h～6 k においては、第 3 の時間単位を「月」とし、第 3 の時間周期階層は、所定数（12 個）の「月」から構成される「年」に係る周期階層となっている。

【0088】

30

このように、特徴量ベクトル生成の際に設定される時間単位及び時間周期階層については、

(ア) 第 1 の時間単位と、所定数の第 1 の時間単位から構成される第 2 の時間単位と、この後、順次 1 つ前の所定数の時間単位から構成される、順次序数を 1 だけ増分させた時間単位とが設定されており、

(イ) 時間周期階層として、第 1 の時間単位に係る時間を周期とする第 1 の時間周期階層と、第 2 の時間単位に係る時間を周期とする第 2 の時間周期階層と、この後、順次序数を 1 だけ増分させた時間単位に係る時間を周期とする時間周期階層とのうちの少なくとも 2 つの時間周期階層が採用されるのである。

【0089】

40

尚、特徴量ベクトル生成の際に設定される時間単位及び時間周期階層の態様は、当然に、上記 (a)～(c) に限定されるものではない。例えば、時間単位として「季」（春夏秋冬）を採用し、対応する時間周期階層を、「季」から構成される「年」とすることもできる。また、時間単位として「半年」を採用し、対応する時間周期階層を、「半年」から構成される「年」や「10 年」とすることも可能である。

【0090】

さらに、上述した具体的な特徴量ベクトルにおける第 1 の時間周期階層「週」に属する同時帯（上記 4. の時間帯）に係る対象データは、6 a、6 b、6 c 及び 6 d の 4 つ（複数）設定されている。これらは、第 1 の時間周期階層「週」を構成する全ての（7 つの）「日」の中の一部の「日」（前日、前々日及び 3 日前の日）に係るデータである。即ち

50

、第1の時間周期階層「週」に係る周期性を予測に反映させるための入力データとして、1週間の中の全ての7日分の対象データを取り扱うのではなく、その一部のデータを採用している。

【0091】

また、同様に、第2の時間周期階層「月」に属する同時帯（上記4.の時間帯）に係る対象データは、6e、6f及び6gの3つ（複数）設定されている。これらは、第2の時間周期階層「月」を構成する全ての「週」の中の一部の「週」（1、2及び3週間前の週）に係るデータである。即ち、第2の時間周期階層「月」に係る周期性を予測に反映させるための入力データとして、1ヶ月の中の全ての4～5週間分の対象データを取り扱うのではなく、その一部のデータを採用している。

10

【0092】

このような対象データの採用をまとめると、時間周期階層の各々に属する時点又は時間間隔に係る複数の対象データ値は、

（ア）第1の時間周期階層を構成する全ての第1の時間単位の中の一部の時間単位に係る対象データ値群と、

（イ）第2の時間周期階層を構成する全ての第2の時間単位の中の一部の時間単位に係る対象データ値群と、この後、

（ウ）順次序数を1だけ増分させた時間周期階層を構成する全ての時間単位の中の一部の時間単位に係る対象データ値群と

のうちの少なくとも2つの群を含むことになっている。その結果、特徴量ベクトルは、時間間隔が不均一であって離散的な時間履歴データから構成された要素を有することになる。

20

【0093】

このように、各時間周期階層に属する対象データの数を離散的に限定することにより、全ての対象データを連続的に採用する場合に比べて、より少ない計算量をもって各時間周期階層に係る周期性を学習させることが可能となる。即ち、連続した全ての時系列データを用いて学習するのではなく、例えば先月や先々月の同時期の離散的なデータをもって学習させることによって、予測のための演算時間の短縮が可能となり、より実用的な処理を実現することができるのである。

【0094】

また、上述した具体的な特徴量ベクトルにおける第1の時間単位「日」、第2の時間単位「週」、及び第3の時間単位「月」に係る対象データ値は、当該時間単位における所定の「時刻」又は「時間帯」（本実施形態では上記4.の時間帯）での値をとっている。このような設定を一般化してまとめると、第1の時間周期階層を構成する全ての第1の時間単位の中の一部の時間単位に係る対象データ値、第2の時間周期階層を構成する全ての第2の時間単位の中の一部の時間単位に係る対象データ値、及び、この後順次序数を1だけ増分させた時間周期階層を構成する全ての時間単位の中の一部の時間単位に係る対象データ値は、自らに係る時間周期階層を構成する時間単位における、第1の時間単位を構成する基本時間単位（本実施形態では「時間帯」）での値をとることとなる。

30

【0095】

このように、各時間周期階層に属する対象データについて、基本時間単位での値に統一することによっても、有効に比較できるデータを離散的に選択することになるので、より少ない計算量をもって各時間周期階層に係る周期性を学習させることが可能となる。

40

【0096】

尚、特徴量ベクトルの要素となる各データ値における時間的な離散度合い（離散間隔）やベクトル要素数 n （次元数）は、予測処理の用途や利用対象に合わせて設定することができる。また、さらなる高精度の予測を実現するために、ベクトル要素の一部として、気象情報（例えば、晴れ、曇り又は雨等の区別）や、交通又は運行情報（例えば、現所在位置での道路渋滞の有無又は程度）を採用することも可能である。

【0097】

50

以上に説明したように、特徴量ベクトルの生成において、複数の時間周期階層の各々について所定時間間隔を有する離散的な対象データ値を採用することによって（特に人間の）行動パターンを効果的に学習し、より精度の高いデータ値の予測を達成することが可能となる。ここで、人間の行動は、所定時間間隔毎に発生する、又は複数の周期性を含むパターンをなすことが多い。例えば、毎週特定の曜日や毎月特定の日に、又は所定の日数の間隔又は所定の月間隔で所定の行動をとる、例えば所定の場所に赴く傾向を有することがよく見られる。従って、上述したような対象データ値からなる特徴量ベクトルを、過去の入力と現在の入力とを相互に影響させるフィードバック結合処理を行うRNNに入力することによって、特徴量ベクトルにおいて配慮されていない時間間隔や周期性に係る行動パターンをも識別し、より精度の高い予測を行うことも可能となるのである。

10

【0098】

ここで、図5（B）を用いて、複数の時間周期階層に係る特徴量ベクトルの要素を決定するための具体的な形態を説明する。

（S51、S52）時系列データに係る時間Hが、予測を行う予測時点（日時）から見て、7日前の時点よりも時間的に近いか否かを判定する。ここで真の判定を行った場合、これら時系列データのうち、予測を行う時点の当日、1日前、2日前及び3日前の（予測時点と）同じ時間帯でのデータ（移動先のデータ）を採用して時系列の特徴量ベクトルを生成する。

【0099】

（S53、S54）一方、ステップS51で偽の判定を行った場合、時系列データに係る時間Hが、予測を行う予測時点（日時）から見て、1ヶ月前の時点よりも時間的に近いか否かを判定する。ここで真の判定を行った場合、これら時系列データのうち、予測を行う時点の1週間前、2週間前及び3週間前の（予測時点と）同じ時間帯でのデータ（移動先のデータ）を採用して時系列の特徴量ベクトルを生成する。

20

（S55、S56）一方、ステップS53で偽の判定を行った場合、時系列データに係る時間Hが、予測を行う予測時点（日時）から見て、12ヶ月（1年）前の時点よりも時間的に近いか否かを判定する。ここで真の判定を行った場合、これら時系列データのうち、予測を行う時点の1ヶ月前、2ヶ月前及び3ヶ月前の（予測時点と）同じ時間帯でのデータ（移動先のデータ）を採用して時系列の特徴量ベクトルを生成する。

【0100】

30

（S57）一方、ステップS55で偽の判定を行った場合、これら時系列データのうち、予測を行う時点の1年前、2年前及び3年前の（予測時点と）同じ時間帯でのデータ（移動先のデータ）を採用して時系列の特徴量ベクトルを生成する。

【0101】

図6及び図7は、図4に示された各ステップの一実施形態を実行するためのプログラムコードを示す模式図である。

【0102】

以下、図6（A）及び（B）並びに図7（A）及び（B）を順次使用して、図4に示された深層学習過程の各ステップを説明するが、ここで、前もってこの深層学習過程の概略を説明する。最初に、学習対象として入力された特徴量が入力層121iから隠れ層121hを介し出力層121cに向かうように入力されることによって学習が行われる。各層では、入力された特徴量データに対し、当該層内のニューロンにおいて結合荷重（重み係数）を乗算して後の層のニューロンに出力する。このような選択的な処理の結果として次の時間に得られる入力を予測学習する仕組みは、一般にフィードフォワード学習と呼ばれている。

40

【0103】

このフィードフォワード学習に対し、バックプロパゲーション学習とは、誤差逆伝播法によって結合荷重を更新していく仕組みである。ここで、誤差逆伝播法とは、出力層121cから出力された値（例えば、移動先の予測値）と、実際に得られた実測値との誤差を、出力層121cから隠れ層121hを介し入力層121iに向けてフィードバックさせ

50

、最小化するように調整する方法である。具体的には、図4に示すように、フィードフォワードでは、データは入力層、隠れ層及び出力層を順次、この方向に移動し、バックプロパゲーションでは、データはフィードフォワードとは逆の方向に移動する。

【0104】

以下、図6及び7を用いて、図4に示された各ステップを具体的に説明する。尚、ステップ1では、予測する時点での状態（即時状態）によって生成された特徴量ベクトルが、入力層121iへ入力される。

【0105】

図6（A）には、ステップ2（図4）のフィードフォワード処理（入力層＋コンテキスト層→隠れ層）を実行するためのプログラムコード例が示されている。ここでは、入力層ニューロンの出力inputs[inp]が、直前の隠れ層ニューロンの複製であるコンテキスト層ニューロンの出力context[con]とともに隠れ層ニューロンに入力される。ここで、入力層ニューロン出力inputs[inp]には、入力層と隠れ層との間の結合荷重（第1の結合荷重）wih[inp][hid]が重み付けされて総和がとられる。さらに、コンテキスト層ニューロン出力context[con]には、コンテキスト層と隠れ層との間の結合荷重（第3の結合荷重）wch[con][hid]が重み付けされて総和がとられる。これらの入力の総和から各ユニット（ニューロン）に設定された閾値を差し引いた値がシグモイド関数（ $f(x)=1/(1+\exp(x))$ ）によって非線形変換され、出力として次のユニット（ニューロン）に伝達される。尚、結合荷重wih[inp][hid]及びwch[con][hid]における学習前の初期値は、例えばランダムな値とすることができる。

【0106】

次いで、図6（B）には、ステップ3（図4）のフィードフォワード処理（隠れ層→出力層）及びステップ4（図4）のコピー処理（隠れ層→コンテキスト層）を実行するためのプログラムコード例が示されている。ここでは、隠れ層ニューロンの出力hidden[hid]に対し、隠れ層と出力層との間の結合荷重（第2の結合荷重）who[hid][out]が重み付けされて総和がとられる。この入力の総和から各ユニット（ニューロン）に設定された閾値を差し引いた値がシグモイド関数によって非線形変換され、出力層ニューロンの出力actual[out]が生成される。尚、結合荷重who[hid][out]における学習前の初期値も、例えばランダムな値とすることができる。

【0107】

さらに、隠れ層ニューロンの出力hidden[hid]は、上述したように出力層ニューロンに出力されつつ、一方でコンテキスト層ニューロンの出力context[con]として複製されている。

【0108】

次いで、図7（A）には、ステップ5（図4）の誤差算出処理を実行するためのプログラムコード例が示されている。ここでは、入力をフィードフォワードしたステップ1～4によって導出された出力actual[out]（予測値）と、実測値（正解）target[out]とを比較して、誤差erro[out]を算出している。ここで、例えば、自動車（ユーザ）の移動先予測の場合、実測値target[out]として、出発の際にユーザによって入力設定された設定目的地の値とすることもできる。また、エンジンを停止した際の所在位置（到着位置）を測位部によって導出した際の値としてもよい。

【0109】

次いで、図7（B）には、ステップ6及び7（図4）のバックプロパゲーション処理を実行するためのプログラムコード例が示されている。ここでは、ステップ5で取得された誤差erro[out]に対しバックプロパゲーション処理を行って学習を実行する。具体的には、予測値actual[out]と正解target[out]とから算出される誤差erro[out]を、出力層から入力層へと遡らせ、より小さくするように伝播させていき、各層の（ネットワーク全体の）結合荷重who[hid][out]、wih[inp][hid]及びwch[inp][hid]を更新していく。このバックプロパゲーション処理は繰り返し実行され、この更新も繰り返し行われる。

【0110】

尚、このように誤差erro[out]のみに対して再学習を実行することによって計算量を大幅に低減することができるのである。

【0111】

ここで、バックプロパゲーション処理では、最急降下法によって誤差の極小値の探索を行っているが、一般に、学習係数（図7（B）のlearnRate）を大きく設定するほど、学習速度はより高くなる。しかしながら、予測値と正解との誤差を結合荷重の関数として見た場合、学習係数が大きすぎると、誤差が増大したり振動したりしてしまう。一方で、学習係数を小さく設定するほどこの振動等の問題は生じにくくなるが、修正量が小さくなって学習速度がより低下してしまう。このように、学習係数を如何に設定するかによって学習速度、即ち収束の速さや、予測精度が大きな影響を受けるのである。これに対し、本発明は、予測を行う予測時点から見て、比較的に近い時系列データと遠い時系列データとを組み合わせ入力し、これらの時系列データが学習に対して異なる形で作用する現象を利用して、学習速度と予測精度とのトレードオフを改善しているのである。

10

【0112】

また、この本発明に係るバックプロパゲーション処理においては、テストデータを用いて時間周期階層毎に学習係数の最適値を模索し決定することによって、さらに予測精度の向上をも可能とするのである。

【0113】

さらに、上述したステップ1～7の実施形態は、RNN特有のコンテキスト層を一次遅れの中間層として複製（写像）することによって入力層と出力層との両方の情報を予測に反映させ、過去の履歴をも網羅して対象データのパターンを識別する。その結果、例えば欠損を含む未知のパターンをも正確に認識することができるようになるのである。

20

【0114】

また、上述したステップ1～7の実施形態では、例えば、出力層から出力された予測値である予測した移動先を、新たな入力データとして入力層に入力し、再度学習を実行することによって、移動先を予測した時点の更に先の時点での移動先を予測することができる。同様に、この再入力の過程を繰り返すことにより、所定時間内における連続した移動先、例えば移動経路をも予測することが可能となる。

【0115】

〔データ予測の実施例〕

図8は、本発明によるデータ予測方法の一実施例及び比較例を示すテーブルである。

30

【0116】

図8に示すように、本実施例を含む実験では、本実施例の効果（予測精度及び学習時間）を検証するため、実験用の履歴データに対して従来方法（時間周期階層を導入していない時系列データを用いた方法）を採用したデータ予測処理を併せて実施し、比較例1～3とした。本実施例及び比較例1～3では、データ予測装置としてのパーソナルコンピュータ（PC）4として、PC端末であるMacBook Pro Mid 2012（2.3GHz Intel Core i7内蔵）を使用した。

【0117】

本実施例及び比較例1～3において、学習対象は所定の3年間における1231のレコードであり、学習回数は20000回であった。特徴量ベクトルの次元数、即ち入力層のニューロン（ユニット）数は、本実施例では24であり、比較例1～3ではそれぞれ24、48及び96であった。また、隠れ層のニューロン（ユニット）数は、本実施例では16であり、比較例1～3ではそれぞれ16、32及び64であった。さらに、学習係数（learnRate）は、本実施例での全ての時間周期階層において0.2に設定され、比較例1～3でも0.2に設定された。また、誤推定率を評価するための入力には、直近の100件のレコードを利用した。

40

【0118】

図8のテーブルに示したように、本実施例では、入力層のニューロン数（特徴量次元数）が4倍である比較例3と同等の、非常に低い誤推定率（4%）を実現できている。しか

50

も、本実施例の学習時間（76.1cpu sec）は、この比較例3の約15分の1と非常に短い。これは、本実施例と同じ入力層ニューロン数（特徴量次元数）を有するものの誤推定率が33%にまで悪化している比較例1と同等の高い学習速度である。このことから、本実施例は、従来と比較して、より高い予測精度とより高い学習速度とをともに実現していることが理解される。

【0119】

以上、詳細に説明したように、本発明によれば、互いに異なる「少なくとも2つの時間周期階層」の各々に係る対象データ値を特徴量要素とする特徴量ベクトルを採用しているので、対象データの有する様々な周期性及びその変化を、効率良く学習することができる。これにより、相当の予測精度を確保しつつも、従来に比べて学習処理のための演算量をより低減することが可能となるのである。

10

【0120】

また、本発明によるデータ予測処理は、当然に、自動車内に持ち込まれたスマートフォンやカーナビゲーション装置での実行に限定されるものではない。例えば、スマートフォンのユーザが徒歩や、自転車、電車等、自動車以外の手段で移動する場合でも、その移動先を予測するのに適用可能である。さらに、本データ予測処理の予測対象も、当然に、移動先に限定されるものではなく、ユーザの期待する又は所望のニュース等の情報プッシュや、為替・株価の予測等の分野での利用も可能である。いずれにしても、本発明の特徴である異なる時間周期階層に基づく特徴量を取り扱う仕組みは、時系列データにかかわる様々な分野におけるデータ予測に適用・拡張可能となっている。

20

【0121】

前述した本発明の種々の実施形態について、本発明の技術思想及び見地の範囲の種々の変更、修正及び省略は、当業者によれば容易に行うことができる。前述の説明はあくまで例であって、何ら制約しようとするものではない。本発明は、特許請求の範囲及びその均等物として限定するものにのみ制約される。

【符号の説明】

【0122】

- 1、3 スマートフォン（データ予測装置）
- 101、201、301 通信インタフェース部
- 102、302 測位部
- 103 タッチパネル・ディスプレイ（TP／DP）
- 104 スピーカ
- 105 マイク
- 106、206 ログ記憶部
- 107、207 結合荷重記憶部
- 108 コンテンツ蓄積部
- 111、311 位置情報取得部
- 112 状況取得部
- 113、313 特徴量生成部
- 114、314 予測結果処理部
- 115、315 ユーザインタフェース（IF）制御部
- 115c 対話処理部
- 12、22 深層学習部
- 121 入力・隠れ層部
- 121i 入力層
- 121h 隠れ層
- 121c コンテキスト層
- 122 出力層部
- 122o 出力層
- 2 データ予測サーバ（データ予測装置）

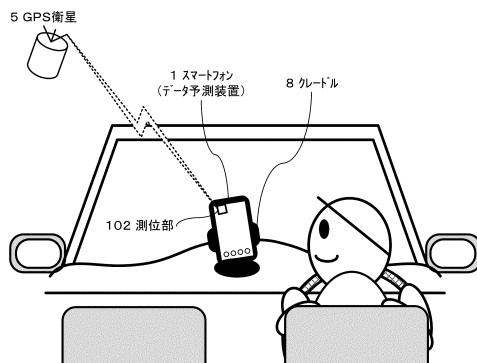
30

40

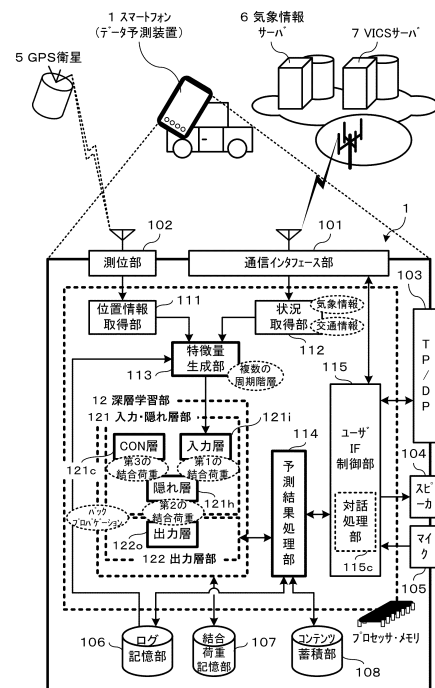
50

- 2 1 1、3 1 2 通信制御部
- 3 0 3 ユーザ I F
- 4 P C (データ予測装置)
- 5 G P S 衛星
- 6 気象情報サーバ
- 7 V I C S (登録商標) サーバ
- 8 クレードル

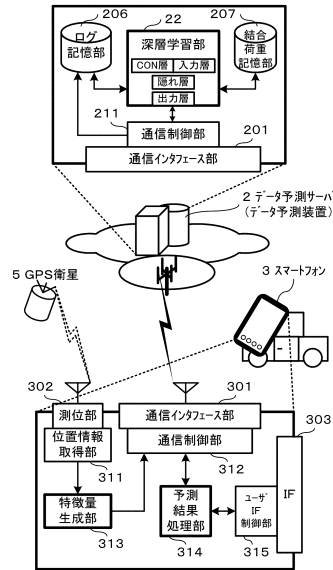
【図 1】



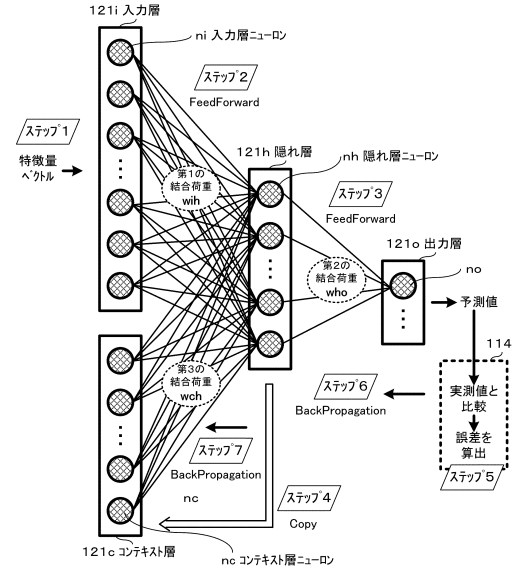
【図 2】



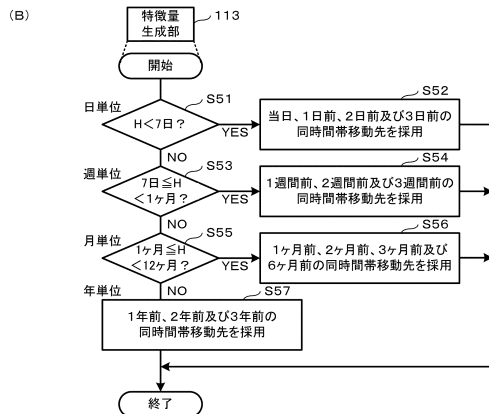
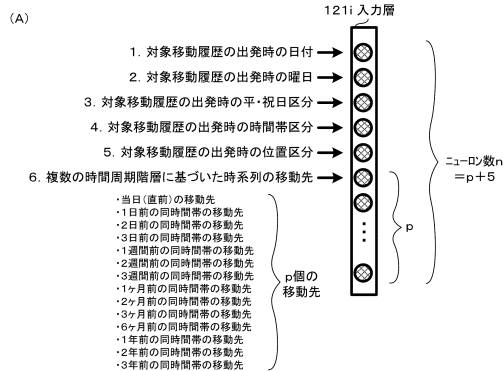
【図 3】



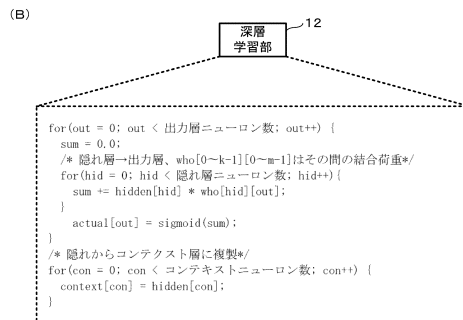
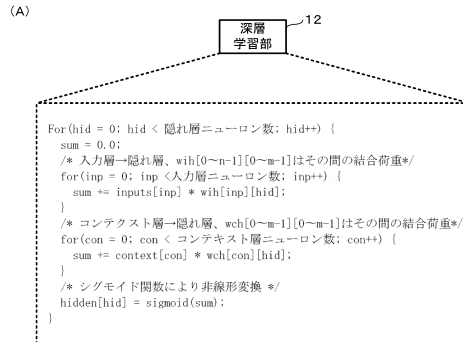
【図 4】



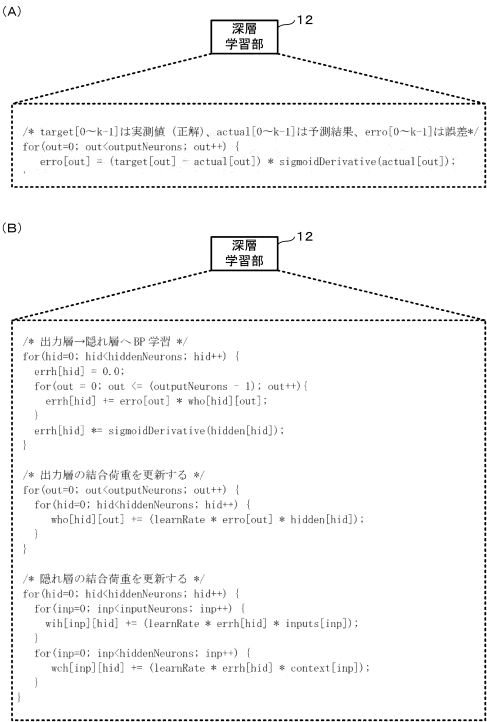
【図 5】



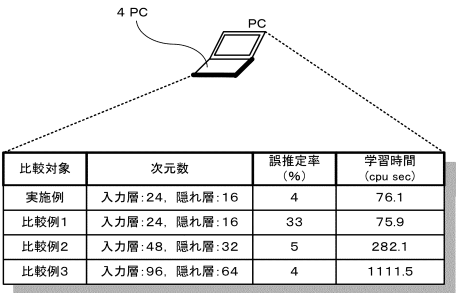
【図 6】



【図 7】



【図 8】



フロントページの続き

(72)発明者 矢崎 智基

埼玉県ふじみ野市大原二丁目1番15号 株式会社KDDI研究所内

審査官 塚田 肇

(56)参考文献 特開平11-102351 (JP, A)

特開平7-319509 (JP, A)

(58)調査した分野(Int.Cl., DB名)

G06N 3/04

G06N 3/08