

11-18-2020

Systems and methods for obtaining optimal mother wavelets to facilitate machine learning tasks

Ishan Sahu

Banerjee

Tanushyam Chattopadhyay

Pal

Utpal Garain

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第6787981号
(P6787981)

(45) 発行日 令和2年11月18日(2020.11.18)

(24) 登録日 令和2年11月2日(2020.11.2)

(51) Int. Cl. F I
 G O 6 N 20/00 (2019.01) G O 6 N 20/00 1 3 0
 G O 6 N 7/00 (2006.01) G O 6 N 7/00

請求項の数 9 外国語出願 (全 17 頁)

<p>(21) 出願番号 特願2018-231781 (P2018-231781) (22) 出願日 平成30年12月11日(2018.12.11) (65) 公開番号 特開2019-121376 (P2019-121376A) (43) 公開日 令和1年7月22日(2019.7.22) 審査請求日 平成31年2月21日(2019.2.21) (31) 優先権主張番号 201721047125 (32) 優先日 平成29年12月28日(2017.12.28) (33) 優先権主張国・地域又は機関 インド(IN)</p>	<p>(73) 特許権者 510337621 タタ コンサルタンシー サービスズ リ ミテッド TATA Consultancy Se rvices Limited インド国 マハーラシュトラ、ムンバイ 400021、ナリマン ポイント、ナー マル ビルディング 9階 Nirmal Building, 9th Floor, Nariman Poin t, Mumbai 400021, Mah arashtra, India. (74) 代理人 110000855 特許業務法人浅村特許事務所</p> <p style="text-align: right;">最終頁に続く</p>
--	---

(54) 【発明の名称】 機械学習タスクを容易にするための最適なマザーウェーブレットを得るためのシステムおよび方法

(57) 【特許請求の範囲】

【請求項1】

機械学習タスクを容易にするための最適なマザーウェーブレットを得るための方法であって、プロセッサに実装された、

複数の信号に基づいて、1つ以上のハードウェアプロセッサによって信号データの第1のセットを識別するステップであって、前記信号データの第1のセットが、1つ以上の信号クラスに対応するラベル付きデータセットを含む、ステップと、

前記信号データの第1のセットおよびマザーウェーブレットの可能性のあるセットに基づいて、前記マザーウェーブレットの可能性のあるセットによるウェーブレット変換に対応する信号の第1のセットのエントロピー比およびマルチレベルウェーブレットエントロピーを含むエネルギー値およびエントロピー値を計算するステップと、

前記エネルギー値およびエントロピー値に基づいて、前記1つ以上の信号クラスに対応する前記信号の第1のセットの重心の値および標準偏差の値を含む信号データの第2のセットを計算するステップと、

前記信号データの第2のセットに基づいて、前記1つ以上の信号クラスに対応する前記重心間の距離を含む距離値のセットを計算するステップであって、前記距離値のセットが、前記信号データの第2のセットに対応する関数に基づいて計算される、ステップと、

前記1つ以上の信号クラスに対応する前記距離値のセットおよび標準偏差の前記値に基づいて、前記信号の第1のセットの前記ウェーブレット変換を実行するための1つ以上の最適なマザーウェーブレットを得るステップと、

を含む、方法。

【請求項 2】

前記距離値のセットを計算するステップが、前記 1 つ以上の最適なマザーウェーブレットを得るために、前記エネルギー値およびエントロピー値に基づいて、N ノルム技法を使用して、前記距離値のセットを正規化するステップを含む、請求項 1 に記載の方法。

【請求項 3】

前記 1 つ以上の最適なマザーウェーブレットを得るステップが、前記 1 つ以上の信号クラスを分類するために、前記ラベル付きデータセットに基づいて前記マザーウェーブレットの可能性のあるセットを解析するステップを含む、請求項 1 に記載の方法。

【請求項 4】

機械学習タスクを容易にするための最適なマザーウェーブレットを得るためのシステムであって、

命令を記憶するメモリと、

1 つ以上の通信インタフェースと、

前記 1 つ以上の通信インタフェースを介して前記メモリに結合された 1 つ以上のハードウェアプロセッサと、

を含み、前記 1 つ以上のハードウェアプロセッサが、前記命令によって、

複数の信号に基づいて、前記 1 つ以上のハードウェアプロセッサによって、1 つ以上の信号クラスに対応するラベル付きデータセットを含む信号データの第 1 のセットを識別し

、

マザーウェーブレットの可能性のあるセットおよび前記信号データの第 1 のセットに基づいて、前記マザーウェーブレットの可能性のあるセットによるウェーブレット変換に対応する信号の第 1 のセットのエントロピー比およびマルチレベルウェーブレットエントロピーを含むエネルギー値およびエントロピー値を計算し、

前記エネルギー値およびエントロピー値に基づいて、前記 1 つ以上の信号クラスに対応する前記信号の第 1 のセットの重心の値および標準偏差の値を含む信号データの第 2 のセットを計算し、

前記信号データの第 2 のセットに基づいて、前記 1 つ以上の信号クラスに対応する前記重心間の距離を含む、前記信号データの第 2 のセットに対応する関数に基づいて計算される距離値のセットを計算し、

前記 1 つ以上の信号クラスに対応する前記距離値のセットおよび前記標準偏差の値に基づいて、前記信号の第 1 のセットのウェーブレット変換を実行するための 1 つ以上の最適なマザーウェーブレットを得る、
ように構成される、システム。

【請求項 5】

前記距離値のセットが、前記 1 つ以上の最適なマザーウェーブレットを得るために、前記エネルギー値およびエントロピー値に基づいて、N ノルム技法を使用して、前記距離値のセットを正規化することによって計算される、請求項 4 に記載のシステム。

【請求項 6】

前記 1 つ以上の最適なマザーウェーブレットが、前記 1 つ以上の信号クラスを分類するために、前記ラベル付きデータセットに基づいて前記マザーウェーブレットの可能性のあるセットを解析することによって得られる、請求項 4 に記載のシステム。

【請求項 7】

1 つ以上のハードウェアプロセッサによって実行されると、前記 1 つ以上のハードウェアプロセッサに機械学習タスクを容易にするための最適なマザーウェーブレットを得るための方法を実行させる 1 つ以上の命令を含む 1 つ以上の非一時的機械可読情報記憶媒体であって、前記方法が、

* (i) 複数の信号に基づいて、前記 1 つ以上のハードウェアプロセッサによって、1 つ以上の信号クラスに対応するラベル付きデータセットを含む信号データの第 1 のセットを識別するステップと、

10

20

30

40

50

(i i) 前記信号データの第 1 のセットおよびマザーウェーブレットの可能性のあるセットに基づいて、前記マザーウェーブレットの可能性のあるセットによるウェーブレット変換に対応する信号の第 1 のセットのエントロピー比およびマルチレベルウェーブレットエントロピーを含むエネルギー値およびエントロピー値を計算するステップと、

(i i i) 前記エネルギー値およびエントロピー値に基づいて、前記 1 つ以上の信号クラスに対応する前記信号の第 1 のセットの重心の値および標準偏差の値を含む信号データの第 2 のセットを計算するステップと、

(i v) 前記信号データの第 2 のセットに基づいて、前記 1 つ以上の信号クラスに対応する前記重心間の距離を含む、前記信号データの第 2 のセットに対応する関数に基づいて計算される距離値のセットを計算するステップと、

(v) 前記 1 つ以上の信号クラスに対応する前記距離値のセットおよび標準偏差の前記値に基づいて、前記信号の第 1 のセットのウェーブレット変換を実行するための 1 つ以上の最適なマザーウェーブレットを得るステップと、
を含む、1 つ以上の非一時的機械可読情報記憶媒体。

【請求項 8】

前記距離値のセットを計算するステップが、前記 1 つ以上の最適なマザーウェーブレットを得るために、前記エネルギー値およびエントロピー値に基づいて、N ノルム技法を使用して、前記距離値のセットを正規化するステップを含む、請求項 7 に記載の 1 つ以上の非一時的機械可読情報記憶媒体。

【請求項 9】

前記 1 つ以上の最適なマザーウェーブレットを得るステップが、前記 1 つ以上の信号クラスを分類するために、前記ラベル付きデータセットに基づいて前記マザーウェーブレットの可能性のあるセットを解析するステップを含む、請求項 7 に記載の 1 つ以上の非一時的機械可読情報記憶媒体。

【発明の詳細な説明】

【技術分野】

【0001】

関連出願および優先権の相互参照

この特許出願は、2017年12月28日にインドで出願されたインド特許出願第201721047125号(名称:機械学習タスクを容易にするための最適なマザーウェーブレットを得るためのシステムおよび方法)の優先権を主張する。

【0002】

本開示は、一般に、機械学習タスクを容易にするための最適なマザーウェーブレットを得ることに関する。より具体的には、本開示は、機械学習タスクを容易にするための最適なマザーウェーブレットを得るためのシステムおよび方法に関する。

【背景技術】

【0003】

一般に、ウェーブレットは、各成分をそのスケールと一致する解像度で解析するために本質的にデータを複数の周波数成分に分割する数学関数である。ウェーブレット解析手順は、解析ウェーブレットまたはマザーウェーブレットと呼ばれるウェーブレットプロトタイプ関数を採用するために実行される。理想的には、ウェーブレットは、複数のタイムシフトと周波数スケールを使用して、時間領域および周波数領域の両方で高解像度で信号を解析するために使用される。ウェーブレットが識別された信号の形態と一致するときはいつでも、離散ウェーブレット変換のようなウェーブレット変換の適用は、信号解析および信号からの特徴抽出を支援する。

【0004】

マザーウェーブレットは、ウェーブレット変換において所与の信号を解析するための基礎を形成する。マザーウェーブレットは、信号クラスに対して実行される特定のタスクに基づいて、アプリケーションごとに異なる場合がある。ウェーブレット変換を適用することによって得られる結果が、選択されたマザーウェーブレットによって影響を受ける可能

10

20

30

40

50

性があるので、最適なマザーウェーブレットの選択は重要である。したがって、信号とマザーウェーブレットとの間にはある程度の相関がなければならない。さらに、多くのマザーウェーブレット関数から選択することができる。したがって、ウェーブレット解析を適用する上での最大の課題は、所与のタスクに対して所望のマザーウェーブレットを選択する方法である。ウェーブレット変換の適用の成功は、タスクのための適切なウェーブレット関数の選択にかかっていると一般的に認められている。

【0005】

従来のシステムおよび方法は、よく知られている技術、例えば連続ウェーブレット変換(CWT)を使用するマザーウェーブレットの選択を提供するが、それらのうちで機械学習タスクに対して実行され得る最適または最良のマザーウェーブレットの選択を提供するものはない。

【発明の概要】

【0006】

以下は、実施形態の基本的な理解を提供するために、本開示のいくつかの実施形態の簡略化された概要を提示する。この概要は、実施形態の広範な概要ではない。実施形態の基本的な/重要な要素を特定すること、または実施形態の範囲を描写することは意図されていない。その唯一の目的は、以下に提示するより詳細な説明の前置きとして、いくつかの実施形態を簡略化された形で提示することである。

【0007】

本開示のシステムおよび方法は、機械学習タスクを容易にするための最適なマザーウェーブレットを得ることを可能にする。本開示の一実施形態では、機械学習タスクを容易にするための最適なマザーウェーブレットを得るための方法が提供され、方法は、複数の信号に基づいて、1つ以上のハードウェアプロセッサによって信号データの第1のセットを識別するステップであって、信号データの第1のセットが、1つ以上の信号クラスに対応するラベル付きデータセットを含む、ステップと、信号データの第1のセットおよびマザーウェーブレットの可能性のあるセットに基づいて、マザーウェーブレットの可能性のあるセットによるウェーブレット変換に対応する信号の第1のセットのエントロピーおよびマルチレベルウェーブレットエントロピーを含むエネルギー値およびエントロピー値を計算するステップと、エネルギー値およびエントロピー値に基づいて、1つ以上の信号クラスに対応する信号の第1のセットの重心の値および標準偏差の値を含む信号データの第2のセットを計算するステップと、信号データの第2のセットに基づいて、1つ以上の信号クラスに対応する重心間の距離を含む距離値のセットを計算するステップであって、距離値のセットが、信号データの第2のセットに対応する関数に基づいて計算される、ステップと、1つ以上の信号クラスに対応する距離値のセットおよび標準偏差の値に基づいて、信号の第1のセットのウェーブレット変換を実行するための1つ以上の最適なマザーウェーブレットを得るステップと、1つ以上の最適なマザーウェーブレットを得るために、エネルギー値およびエントロピー値に基づいて、Nノルム技法を使用して、距離値のセットを正規化することによって距離値のセットを計算するステップと、1つ以上の信号クラスを分類するために、ラベル付きデータセットに基づいて、マザーウェーブレットの可能性のあるセットを解析することによって1つ以上の最適なマザーウェーブレットを得るステップと、得られた1つ以上の最適なマザーウェーブレットによって、信号の第2のセットのウェーブレット変換を実行することによって、信号の第2のセットの分解を可能にするステップを含むステップであって、信号の第2のセットが、信号の第2のセットに対応する1つ以上の信号クラスの分類または回帰によって機械学習タスクを容易にするステップとを含む。

【0008】

本開示の一実施形態では、機械学習タスクを容易にするための最適なマザーウェーブレットを得るためのシステムが提供され、システムは、1つ以上のプロセッサと、1つ以上のプロセッサに動作可能に結合され、1つ以上のプロセッサによる実行のために構成された命令を格納するように構成された1つ以上のデータ記憶装置とを含み、命令は、複数の

10

20

30

40

50

信号に基づいて、1つ以上のハードウェアプロセッサによって、1つ以上の信号クラスに対応するラベル付きデータセットを含む信号データの第1のセットを識別することと、マザーウェーブレットの可能性のあるセットおよび信号データの第1のセットに基づいて、マザーウェーブレットの可能性のあるセットによるウェーブレット変換に対応する信号の第1のセットのエントロピー比およびマルチレベルウェーブレットエントロピーを含むエネルギー値およびエントロピー値を計算することと、エネルギー値およびエントロピー値に基づいて、1つ以上の信号クラスに対応する信号の第1のセットの重心の値および標準偏差の値を含む信号データの第2のセットを計算することと、信号データの第2のセットに基づいて、1つ以上の信号クラスに対応する重心間の距離を含む、信号データの第2のセットに対応する関数に基づいて計算される距離値のセットを計算することと、1つ以上の信号クラスに対応する距離値のセットおよび標準偏差の値に基づいて、信号の第1のセットのウェーブレット変換を実行するための1つ以上の最適なマザーウェーブレットを得ることと、1つ以上の最適なマザーウェーブレットを得るために、エネルギー値およびエントロピー値に基づいて、Nノルム技法を使用して、距離値のセットを正規化することと、1つ以上の信号クラスを分類するために、ラベル付きデータセットに基づいて、マザーウェーブレットの可能性のあるセットを解析することと、得られた1つ以上の最適なマザーウェーブレットによって、信号の第2のセットのウェーブレット変換を実行することによって、信号の第2のセットの分解を可能にすることを含むことと、信号の第2のセットが、信号の第2のセットに対応する1つ以上の信号クラスの分類または回帰によって機械学習タスクを容易にする含むこととを含む。

10

20

【0009】

本開示の一実施形態では、1つ以上のハードウェアプロセッサによって実行されたときに1つ以上のハードウェアプロセッサに機械学習タスクを容易にするための最適なマザーウェーブレットを得るための方法を実行させる1つ以上の命令を含む1つ以上の非一時的機械可読情報記憶媒体が提供され、方法は、複数の信号に基づいて、1つ以上のハードウェアプロセッサによって、1つ以上の信号クラスに対応するラベル付きデータセットを含む信号データの第1のセットを識別するステップと、信号データの第1のセットおよびマザーウェーブレットの可能性のあるセットに基づいて、マザーウェーブレットの可能性のあるセットによるウェーブレット変換に対応する信号の第1のセットのエントロピー比およびマルチレベルウェーブレットエントロピーを含むエネルギー値およびエントロピー値を計算するステップと、エネルギー値およびエントロピー値に基づいて、1つ以上の信号クラスに対応する信号の第1のセットの重心の値および標準偏差の値を含む信号データの第2のセットを計算するステップと、信号データの第2のセットに基づいて、1つ以上の信号クラスに対応する重心間の距離を含む、信号データの第2のセットに対応する関数に基づいて計算される距離値のセットを計算するステップと、1つ以上の信号クラスに対応する距離値のセットおよび標準偏差の値に基づいて、信号の第1のセットのウェーブレット変換を実行するための1つ以上の最適なマザーウェーブレットを得るステップと、1つ以上の最適なマザーウェーブレットを得るために、エネルギー値およびエントロピー値に基づいて、Nノルム技法を使用して、距離値のセットを正規化することによって距離値のセットを計算するステップと、1つ以上の信号クラスを分類するために、ラベル付きデータセットに基づいて、マザーウェーブレットの可能性のあるセットを解析することによって1つ以上の最適なマザーウェーブレットを得るステップと、得られた1つ以上の最適なマザーウェーブレットによって、信号の第2のセットのウェーブレット変換を実行することによって、信号の第2のセットの分解を可能にするステップを含むステップと、信号の第2のセットが、信号の第2のセットに対応する1つ以上の信号クラスの分類または回帰によって機械学習タスクを容易にするステップとを含む。

30

40

【0010】

前述の一般的な説明および以下の詳細な説明の両方は、例示的および説明的なものに過ぎず、本開示を請求するとおりに限定するものではないことを理解されたい。

【0011】

50

本明細書における実施形態は、図面を参照する以下の詳細な説明からよりよく理解されるであろう。

【図面の簡単な説明】

【0012】

【図1】本開示の一実施形態による、機械学習タスクを容易にするための最適なマザーウェーブレットを得るためのシステムのブロック図を示す。

【図2】本開示の一実施形態による、機械学習タスクを容易にするための最適なマザーウェーブレットの取得に含まれるステップを示すフローチャートである。

【図3】本開示の一実施形態による、最適なマザーウェーブレットを得るための1つ以上の信号クラスの中の1つのクラスに対応する複数の信号のグラフ表示を示す。

10

【図4】本開示の一実施形態による、最適なマザーウェーブレットを得るための1つ以上の信号クラスの中の別のクラスに対応する複数の信号のグラフ表示を示す。

【図5】本開示の一実施形態によって解析された各マザーウェーブレットに対応して正規化された距離値のセットのサンプル結果のセットのグラフ表示を示す。

【発明を実施するための形態】

【0013】

本明細書の実施形態ならびにその様々な特徴および有利な詳細は、添付の図面に示され以下の説明に詳述される非限定的な実施形態を参照してより十分に説明される。本明細書で使用される例は、単に本明細書の実施形態を実施することができる方法の理解を容易にし、当業者が本明細書の実施形態を実施することをさらに可能にすることを意図している。したがって、例は、本明細書の実施形態の範囲を限定するものとして解釈されるべきではない。

20

【0014】

本開示の実施形態は、機械学習タスクを容易にするための最適なマザーウェーブレットを得るためのシステムおよび方法を提供する。ウェーブレット解析は、一般的な時間周波数解析方法であり、例えば生物学的信号、振動信号、音響および超音波信号を含む広範囲の信号を解析するために様々な分野で適用されてきた。時間領域と周波数領域の両方の情報を提供する機能を備えたウェーブレット解析は、主に信号の時間周波数解析、信号圧縮、信号雑音除去、特異点解析、および特徴抽出のためのものである。同じ信号に異なるマザーウェーブレットを適用すると異なる結果が生じる可能性があるため、ウェーブレット変換を実行する際の主な課題は、所与のタスクに最適なマザーウェーブレットを選択することである。

30

【0015】

機械学習技術は、訓練後に入力が入力が正しく分類されるようにアルゴリズムのパラメータを自動的に調整することによって動作する。残念なことに、訓練は、アルゴリズムおよび他の複雑なタスクに対して何十万ものペア（入力、ターゲット）を提示することを含み得る。従来のシステムおよび方法は、履歴信号データを使用して特徴のセットを生成し、生成された特徴のセットに基づいて分類器を訓練し、分類器モデルを抽出し、最後に分類器モデルに基づいてクラスラベル予測を生成することによって信号分類を提供する。

【0016】

40

したがって、ウェーブレット解析技術のいずれかを使用して信号のセットの分解を実行することができるマザーウェーブレットを推奨することを容易にする技術が必要であり、これにより異なるクラスの信号間の区別または分離が明らかになる。さらに、この技術は、信号とクラスラベルとの間のマッピングが自動的に学習され得る機械学習タスクの実行を容易にしなければならず、その結果、そのようなラベルが不明であると考えられる中で信号に対する試験または実世界展開中に予測を行うことができる。要するに、この技術は信号クラス間の明確な区別を提供し、機械学習タスクをさらに容易にしなければならない。

【0017】

ここで同様の参照符号は図面全体を通して一貫して対応する特徴を示す図面、特に図1

50

～図5を参照すると、好ましい実施形態が示されており、これらの実施形態は以下の例示的なシステムおよび/または方法に関連して説明される。

【0018】

図1は、本開示の一実施形態による、機械学習タスクを容易にするための最適なマザーウェーブレットを得るためのシステム100の例示的なブロック図を示す。一実施形態では、システム100は、1つ以上のプロセッサ104、通信インタフェースデバイスまたは入出力(I/O)インタフェース106、および1つ以上のプロセッサ104に動作可能に結合された1つ以上のデータ記憶装置またはメモリ102を含む。ハードウェアプロセッサである1つ以上のプロセッサ104は、1つ以上のマイクロプロセッサ、マイクロコンピュータ、マイクロコントローラ、デジタル信号プロセッサ、中央処理装置、ステートマシン、論理回路、および/または動作命令に基づいて信号を操作する任意のデバイスとして実装することができる。他の機能の中でも、プロセッサは、メモリに格納されているコンピュータ可読命令を取り出して実行するように構成されている。一実施形態では、システム100は、ラップトップコンピュータ、ノートブック、携帯用デバイス、ワークステーション、メインフレームコンピュータ、サーバ、ネットワーククラウドなどの様々なコンピューティングシステムに実装することができる。

10

【0019】

I/Oインタフェースデバイス106は、様々なソフトウェアおよびハードウェアインタフェース、例えばウェブインタフェース、グラフィカルユーザーインタフェースなどを含むことができ、LAN、ケーブルなどのような有線ネットワーク、およびWLAN、セルラー、または衛星などの無線ネットワークを含む広範囲のネットワークN/Wおよびプロトコルタイプ内での複数の通信を容易にすることができる。一実施形態では、I/Oインタフェースデバイスは、複数のデバイスを互いにまたは別のサーバに接続するための1つ以上のポートを含むことができる。

20

【0020】

メモリ102は、例えば、スタティックランダムアクセスメモリ(SRAM)およびダイナミックランダムアクセスメモリ(DRAM)のような揮発性メモリ、および/または読み出し専用メモリ(ROM)、消去可能プログラマブルROM、フラッシュメモリ、ハードディスク、光ディスク、磁気テープなどのような不揮発性メモリを含む当技術分野で既知の任意のコンピュータ可読媒体を含むことができる。

30

【0021】

図1を参照して、図2は、本開示の一実施形態による、機械学習タスクを容易にするための最適なマザーウェーブレットを得るための方法の例示的なフロー図を示す。一実施形態では、システム100は、1つ以上のハードウェアプロセッサ104に動作可能に結合されたメモリ102の1つ以上のデータ記憶装置を含み、1つ以上のプロセッサ104による方法のステップの実行のための命令を格納するように構成される。本開示の方法のステップは、図1およびフロー図に示されるようなシステム100の構成要素を参照して説明される。本開示の実施形態では、ハードウェアプロセッサ104は、命令が構成されると、本明細書に記載の1つ以上の方法を実行する。

【0022】

本開示の一実施形態によれば、ステップ201において、1つ以上のハードウェアプロセッサ104は、複数の信号に基づいて信号データの第1のセットを識別し、信号データの第1のセットは、1つ以上の信号クラスに対応するラベル付きデータセットを含む。一実施形態では、複数の信号は、1つ以上のセンサ(図示せず)を使用して得ることができる。一実施形態では、ラベル付きデータセットを得ることは、ラベルなしデータを入力として得ることと、ラベルなしデータの各部分を何らかの意味のあるまたは知るのが望ましい何らかの意味のある「タグ」、「ラベル」または「クラス」で増強することを含む。図3および図4を参照して、1つ以上のセンサを通して得られた複数の信号(波形)が検討され、図3は、1つのクラス、例えば1つ以上のクラスの中のクラス1に対応する複数の信号を示し、図4は、別のクラス、例えば1つ以上の信号クラスの中のクラス2に対応

40

50

する複数の信号を示す。以下の表 1 を参照すると、複数の信号に基づいてラベル付きデータセットを含む信号データの第 1 のセットが参照されてもよく、各行は 1 つ以上のセンサを通して得られた複数の信号の中のある信号に対応し、ここで、第 1 の列、すなわち列 A は 1 つ以上の信号クラスを示し、列 B、C、D および E は記録の各時間ステップにおける振幅値のセットを示す。

【表 1】

表 1

信号クラス/ ラベル (A)	信号振幅値 (B)	信号振幅値 (C)	信号振幅値 (D)	信号振幅値 (E)
1	0	-0.00085	-0.000153	-0.00212
1	0	-0.00096	-0.00089	-0.00028
1	0	0.004518	0.013311	0.025202
1	0	0.000661	0.002034	0.003762
1	0	-0.00044	-0.00021	0.000653

10

【0023】

本開示の一実施形態によれば、ステップ 202 において、信号データの第 1 のセットおよびマザーウェーブレットの可能性のあるセットに基づいて、エネルギー値およびエントロピー値が計算されてもよく、エネルギー値およびエントロピー値は、マザーウェーブレットの可能性のあるセットによるウェーブレット変換に対応する信号の第 1 のセットのエントロピー比およびマルチレベルウェーブレットエントロピーを含む。一実施形態では、マザーウェーブレットの可能性のあるセットの識別は、任意の従来技法または方法を通じて実行され得る。例えば、1 つ以上のハードウェアプロセッサ 104 は、ウェブおよび出版物をクロールしてウェーブレットキーワードのリストまたは表を抽出することによって、マザーウェーブレットの可能性のあるセットを得ることができる。さらに、識別は、ディープニューラルネットワーク技術（本明細書では説明されない）によって実行されてもよく、ここではマザーウェーブレットの可能性のあるセットを得るために、入力信号が様々なレベルのネットワークを通過する。

20

【表 2】

表 2

マザーウェーブレット (サブセット)
h a a r
d b 1
d b 2
d b 3
s y m 2
s y m 3
c o i f 1
c o i f 2
c o i f 3

30

40

【0024】

この実施形態では、識別されたマザーウェーブレットの可能性のあるセットを使用して、信号データの第 1 のセットにウェーブレット変換を適用することによって信号の第 1 のセットを分解することができる。信号の第 1 のセットの形態に一致する（マザーウェーブレットの可能性のあるセットの中からの）いずれかのマザーウェーブレットが識別されるときはいつでも、ウェーブレット変換の適用は（第 1 のセットの中からの）信号の解析およびそこからの特徴抽出を容易にする。これにより、ウェーブレット変換の全プロセスは

50

、機械学習タスクを容易にするためまたはその他のユースケースの自動化のために、（第1の信号のセットに対応する）信号の1つ以上の係数（第1のレベルの係数および詳細な係数を含む）を得ること、特徴抽出および最終的にはこれらの特徴を使用することを目的とする。一実施形態では、ウェーブレット変換は、新しい関数を得るために以前の関数の値を変更することを可能にする基本関数、またはマザーウェーブレットの可能性のあるセットによって記述することができる。

例示的な実装形態では、1つ以上の係数は、ウェーブレット変換を実行することによって以下のように得ることができる。

近似係数 = - 0 . 0 0 0 5 9 8 9 4 8、 - 0 . 0 0 2 5 8 1 4 3 5、 - 0 . 0 0 3 8 3 3 8 6 2、および

詳細係数 = 0 . 0 0 0 5 9 8 9 4 8、 0 . 0 0 0 4 1 9 9 5 1、 0 . 0 0 0 1 6 9 3 5 2

【 0 0 2 5 】

一実施形態では、（マザーウェーブレットの可能性のあるセットの中からの）単一のマザーウェーブレットが、無限数の子ウェーブレットを表すための基礎を形成する。別の用語には、ウェーブレットファミリー、マザーウェーブレット、および子ウェーブレットが含まれる。ウェーブレット変換を実行する一般的な方法は、連続ウェーブレット変換（CWT）、離散ウェーブレット変換（DWT）、高速ウェーブレット変換（FWT）、リフティングスキームおよび一般化リフティングスキーム、ウェーブレットパケット分解（WPD）、定常ウェーブレット変換（SWT）、分数フーリエ変換（FRFT）および分数ウェーブレット変換（FRWT）を含む。

【 0 0 2 6 】

本開示の一実施形態によれば、1つ以上の係数に基づいて、エネルギー値およびエントロピー値は、例えば、シャノンのエントロピー技術を使用して得ることができる。1つ以上の信号クラスに対応する信号のエネルギー値およびエントロピー値は、値、例えばエネルギー/エントロピーの比率を測定する。信号*i*に対する比率を x_i で表すこととする。したがって、信号データの第1のセットには、（複数の信号のうちの）各信号の1つの測定値がある。さらに、（複数の信号のうちの）信号をマザーウェーブレットの可能性のあるセットでウェーブレット変換を使用して得られた*n*個のウェーブレット係数のエネルギーは、以下の式（1）で表すことができる。

【数1】

$$Energy = \sum_{i=1}^n |C_i|^2 \quad \text{式 (1)}$$

1つ以上の係数に対するシャノンのエントロピーは、以下の式（2）のように定義することができる。

【数2】

$$Entropy(Shannon) = -\sum_{i=1}^n p_i \log p_i \quad \text{式 (2)}$$

ここで、 p_i は、各ウェーブレット係数に対するエネルギー確率の分布 C_i である。

したがって、マザーウェーブレットの可能性のあるセットを使用した（複数の信号のうちの）信号についてのエネルギーとシャノンのエントロピーの比率（ x_i で表される）は、次のように定義することができる。

【数3】

$$P_i = \frac{|C_i|^2}{E_{energy}} \quad \text{式 (3)}$$

10

20

30

40

50

および

【数 4】

$$\sum_{i=1}^n p_i = 1 \quad \text{式 (4)}$$

例示的な実装では、エネルギーとシャノンのエントロピーの比率に対応する1つ以上の係数ならびに式(3)および(4)に基づいて、信号の第1のセットのエントロピーおよびマルチレベルウェーブレットエントロピーを含むエネルギー値およびエントロピー値は、次のように得ることができる。

エネルギー値 = 1 1 2 . 4 8 1 8 1 0 3、
エントロピー値 = - 2 4 . 0 2 9 1 1 5、および
エントロピー比 = - 4 . 6 8 1 0 6 3 3 8

本開示の実施形態は、エネルギー値およびエントロピー値の計算を、エネルギーとシャノンのエントロピーの比率の使用のみに限定しないことに留意されたい。一実施形態では、エネルギー値およびエントロピー値を得るために、同様の他の計算、例えば、 x_i を得るためのマルチレベルウェーブレット時間エネルギーシャノンエントロピー(MWTEE)を実行することができる。

【0027】

本開示の一実施形態によれば、ステップ203において、エネルギー値およびエントロピー値に基づいて、1つ以上の信号クラスに対応する、信号の第1のセットの重心の値および標準偏差の値を含む信号データの第2のセットを計算することができる。一実施形態では、1つ以上の信号クラスに対応する(複数の信号のうちの)各信号に属するすべての x_i を使用して、重心の値および標準偏差の値を計算することができる。

一実施形態では、重心の値は、

【数 5】

$$\frac{x_1 + x_2 + x_3 + \dots + x_m}{m}$$

から得ることができ、

ここで、ベクトル空間 R^n における有限点集合

【数 6】

$$x_1 + x_2 + x_3 \dots + x_m$$

の重心は、

【数 7】

$$\frac{x_1 + x_2 + x_3 + \dots + x_m}{m}$$

である。

一実施形態では、有限点集合の標準偏差は、次のように計算することができる。

【数 8】

$$\sqrt{\frac{1}{m-1} \sum_{i=1}^m |x_i - u|^2}$$

ここで

【数 9】

$$u = \frac{1}{m} \sum_{i=1}^m x_i$$

10

一実施例では、上記の式を使用して、1つ以上の信号クラスに対応する信号データの第2のセットを、以下の表3に示すように得ることができる。

【表 3】

表 3

	クラス 1	クラス 2
重心の値	-0.424363441708168	-1.44449956902769
標準偏差の値	1.58208913535412	6.15306158570416

20

【0028】

本開示の一実施形態によれば、ステップ204において、信号データの第2のセットに基づいて、1つ以上の信号クラスに対応する重心間の距離を含む距離値のセットが計算され、ここで距離値のセットは、信号データの第2のセットに対応する関数に基づいて計算される。一実施形態では、距離値のセットは、重心の値および標準偏差の値に基づいて、任意の既知の技法および方法、例えばユークリッド距離技法を使用して計算することができる。ユークリッド距離技法では、ユークリッド距離またはユークリッド計量は、ユークリッド空間内の2点間の通常の直線距離を含む。この距離によって、ユークリッド空間は計量空間となる。当技術分野で既知のように、計量は、集合上の非負関数 d、すなわち

【数 10】

$$d: X \times X \rightarrow (0, \infty)$$

30

を含む。Xに属する x、y、z については、以下の条件が満たされる。

【数 11】

$$d(x, y) \geq 0$$

40

【数 12】

$$d(x, y) = 0 \iff x = y$$

【数 13】

$$d(x, y) = d(y, x)$$

および

50

【数 1 4】

$$d(x,z) \leq d(x,y) + d(y,z)$$

【0 0 2 9】

例示的な実装形態では、ユークリッド距離技法を使用して、（各マザーウェーブレットの可能性のあるセットについての）重心間の距離を含む距離値のセットは、以下の表 4 に示すように得ることができる。

【表 4】

表 4

(マザーウェーブレットの可能性のあるセットの中からの) マザーウェーブレット	距離値のセット (クラス重心間の距離)	標準偏差の値 (ステップ 2 0 3 で取得)
「h a a r」	3. 7 6 5 1 2 5 2 6 5 8 7 4 8 0	6. 1 0 8 5 4 5 1 1 9 4 3 6 8 2
「d b 1」	3. 7 6 5 1 2 5 2 6 5 8 7 4 8 0	6. 1 0 8 5 4 5 1 1 9 4 3 6 8 2
「d b 2」	3. 6 1 1 1 1 0 9 5 2 8 8 2 2 0	5. 8 5 7 7 4 5 9 8 1 6 7 5 8 9
「d b 3」	3. 6 0 9 4 6 8 0 5 0 0 9 4 7 6	5. 8 5 1 1 9 6 1 4 2 5 4 1 5 1

10

【0 0 3 0】

本開示の一実施形態によれば、距離値のセットは、N ノルム技法または最大距離値技法を使用して正規化することができる。当技術分野で既知のように、ノルムは、ゼロベクトルを除いて、ベクトル空間内の各ベクトルに厳密に正の長さまたはサイズを割り当てる関数であり得る。例えば、n - dimension ユークリッド空間 R^n では、ベクトル

20

【数 1 5】

$$x = (x_1, x_2, \dots, x_n)$$

の長さの直感的な概念は、以下のように計算することができる。

【数 1 6】

$$\|X\|_2 := \sqrt{x_1^2 + \dots + x_n^2}$$

30

例示的な実装形態では、正規化を実行した後に、以下の表 5 に示すように距離値のセットを得ることができる。

【表 5】

表 5

(マザーウェーブレットの可能性のあるセットの中からの) マザーウェーブレット	正規化された距離値
「h a a r」	0. 8 0 7 6 4 6 7 5 1
「d b 1」	0. 8 0 7 6 4 6 7 5 1
「d b 2」	0. 7 7 4 6 0 9 5 6
「d b 3」	0. 7 7 4 2 5 7 1 4 6

40

本開示の実施形態は、距離値のセットの計算を、ユークリッド距離技法のみの使用に限定しないことに留意されたい。実施形態は、その他の技法（例えば、性能計量技法または

50

nボール八面体技法) またはそれらの組み合わせを使用して距離値のセットを計算することを提供する。適切な距離関数の選択は、ラベル付きデータセットの種類およびラベル付きデータセットの対応する特性に基づいてもよい。さらに、本開示の実施形態は、その他の技法またはそれらの組み合わせを使用して距離値のセットの正規化の実行をサポートする。

【0031】

本開示の一実施形態によれば、ステップ205において、距離値のセットおよび1つ以上の信号クラスに対応する標準偏差の値に基づいて、信号の第1のセットのウェーブレット変換を実行するための1つ以上の最適なマザーウェーブレットを得ることができる。このようにして、(上記のステップ204で計算された距離値のセットから)最大距離値に対応する1つ以上のマザーウェーブレットが選択される。したがって、一実施形態では、複数のマザーウェーブレットが最大距離値に対応する場合、標準偏差の最小値を有するマザーウェーブレットを選択することができる。

10

【0032】

図5を参照して、2つの信号クラスを有するMIMIC BPデータセット(Johnson AEW, Pollard TJ, Shen L, Lehman L, Feng M, Ghassemi M, Moody B, Szolovits P, Celi LA、およびMark RG, Scientific Data(2016)による自由にアクセス可能な救命救急データベース)に関するサンプル結果が検討される。上記で解析されたマザーウェーブレットのそれぞれに対応して、プロットされた重心間の計算された正規化距離が参照されてもよい。再び図5を参照すると、(マザーウェーブレットの可能性のあるセットの中から)「db10」マザーウェーブレットが(上記のステップ202~204に示された計算例に基づいて)クラス間の最大の区別を与え、それゆえ最適なマザーウェーブレットとして選択されてもよいことに留意されたい。

20

【0033】

本開示の一実施形態によれば、得られた1つ以上の最適なマザーウェーブレットは、信号の第2のセットの離散ウェーブレット変換を実行することによって信号の第2のセットの分解を可能にし、ここで信号の第2のセットは、信号の第2のセットに対応する1つ以上の信号クラスの分類または回帰によって機械学習実装を容易にする。例えば、上記のステップ205を参照すると、「db10」が最適なマザーウェーブレットとして得られる場合、「db10」は、信号の第2のセットの中から同じタイプの同様にラベル付けされた信号(新しいセットの信号)の分類のような機械学習タスクにおいて実装され得る。従来のシステムおよび方法は、履歴信号データを使用して特徴のセットを生成し、生成された特徴のセットに基づいて履歴信号データの訓練部分で分類器を訓練し、分類器モデルを抽出し、最後に分類器モデルに基づいてクラスラベル予測を生成することによって信号分類を提供する。

30

【0034】

本開示の実施形態は、履歴信号データから特徴のセットを生成することを容易にする。ウェーブレット変換に基づいて特徴のセットを生成する間、マザーウェーブレットは最適でなければならない。最適なマザーウェーブレットの選択は、より高品質な特徴の生成を容易にする。選択された最適なマザーウェーブレットは、異なるクラスの信号間の区別を容易にする特徴のセットを提供する。これは学習を容易にし、良好な分類器または回帰モデルおよび第2の信号のセットに関する予測をもたらす。一実施形態では、得られた1つ以上の最適なマザーウェーブレットに基づいてウェーブレット変換を実行した後に計算することができるいくつかの特徴は、歪度、尖度、エネルギー、ゼロ交差率、平均、分散などを含む。ウェーブレット変換に基づくこれらすべての特徴の計算は、1つ以上の最適なマザーウェーブレットに依存する。

40

【0035】

一実施形態では、メモリ102は、機械学習タスクを容易にするための最適なマザーウェーブレットを得ることに関連する任意のデータを格納するように構成することができる

50

。一実施形態では、信号データの第1のセット、マザーウェーブレットの可能性のあるセット、重心の値および標準偏差の値、信号データの第2のセット、ならびに距離値のセットなどに関する情報は、メモリ102に格納される。さらに、機械学習タスクを容易にするための最適なマザーウェーブレットを得ることに關するすべての情報（入力、出力など）は、参照のために履歴データとしてデータベースにも格納することができる。

【0036】

記述された説明は、当業者が実施形態を作成し使用することを可能にする本明細書の主題を説明する。主題の実施形態の範囲は、特許請求の範囲によって定義され、当業者が思いつく他の修正を含むことができる。そのような他の修正は、それらが特許請求の範囲の文字通りの言葉と異なる類似の要素を有する場合、またはそれらが特許請求の範囲の文字通りの言葉との差異が薄弱である同等の要素を含む場合、特許請求の範囲内にあることが意図されている。

10

【0037】

保護の範囲は、そのようなプログラム、さらにメッセージを有するコンピュータ可読手段に拡張されることが理解されるべきであり、そのようなコンピュータ可読記憶手段は、プログラムがサーバまたはモバイル装置または任意の適切なプログラブル装置上で動作する場合、この方法の1つ以上のステップを実装するためのプログラムコード手段を含む。ハードウェア装置は、サーバまたはパーソナルコンピュータなどのような任意の種類のコンピュータ、あるいはそれらの任意の組み合わせなど、プログラム可能な任意の種類の装置とすることができる。装置はまた、例えば特定用途向け集積回路（ASIC）、フィールドプログラブルゲートアレイ（FPGA）などのハードウェア手段、または例えばASICとFPGAといったハードウェア手段とソフトウェア手段との組み合わせ、または少なくとも1つのマイクロプロセッサと中にソフトウェアモジュールが配置された少なくとも1つのメモリなどであり得る手段を含むことができる。したがって、手段は、ハードウェア手段およびソフトウェア手段の両方を含むことができる。本明細書に記載の方法の実施形態は、ハードウェアおよびソフトウェアで実装することができる。装置はソフトウェア手段も含むことができる。代替的に、実施形態は、複数のCPUを使用して、異なるハードウェア装置上で実装されてもよい。

20

【0038】

本明細書の実施形態は、ハードウェア要素およびソフトウェア要素を含むことができる。ソフトウェアで実装される実施形態は、ファームウェア、常駐ソフトウェア、マイクロコードなどを含むがこれらに限定されない。本明細書で説明される様々なモジュールによって実行される機能は、他のモジュールまたは他のモジュールの組み合わせで実装されてもよい。この説明の目的のために、コンピュータ使用可能またはコンピュータ可読媒体は、命令実行システム、機器、または装置によって使用するための、またはこれらに關係するプログラムを含み、格納し、通信し、伝播し、または移送することができる任意の機器とすることができる。

30

【0039】

示されたステップは、示された例示的な実施形態を説明するために提示されており、進行中の技術開発によって特定の機能が実行される方法が変更されることが予期されるべきである。これらの例は、本発明を説明するためのものであって、限定するものではない。さらに、機能的成分の境界は、説明の便宜上、本明細書では任意に定義されている。指定された機能およびそれらの關係が適切に実行される限り、代替の境界を定義することができる。本明細書に含まれる教示に基づいて、当業者には代替物（本明細書に記載されているものの均等物、拡張物、変形物、逸脱物などを含む）が明らかであろう。そのような代替物は、開示された実施形態の範囲および精神の範囲内に入る。また、「含む」、「有する」、「包含する」、および「含有する」ならびに他の同様の形態の単語は、等価な意味を有するように意図され、これらの単語のいずれかに続く項目は、そのような項目の網羅的なリストであることを意味しない、またはリストされた項目のみに限定されることを意味する点で非限定的であることが意図されている。本明細書および添付の特許請求の範囲

40

50

で使用する単数形「a」、「an」および「the」は、文脈上他に明確に指示しない限り、複数形の参照を含むことにも留意されたい。

【0040】

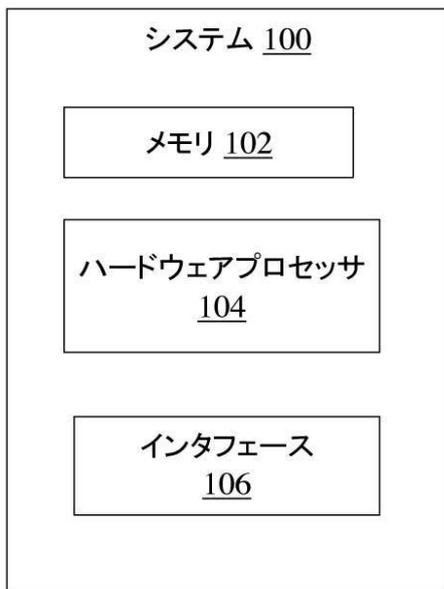
さらに、本開示に即した実施形態を実装する際に、1つ以上のコンピュータ可読記憶媒体を利用することができる。コンピュータ可読記憶媒体は、プロセッサによって読み取り可能な情報またはデータを格納することができる任意のタイプの物理メモリを指す。したがって、コンピュータ可読記憶媒体は、プロセッサに本明細書に記載の実施形態に即したステップまたはステージを実行させるための命令を含む、1つ以上のプロセッサによる実行のための命令を格納することができる。「コンピュータ可読媒体」という用語は、有形の品目を含み、搬送波および過渡信号、すなわち非一時的なものは除外されると理解すべきである。例として、ランダムアクセスメモリ(RAM)、読出し専用メモリ(ROM)、揮発性メモリ、不揮発性メモリ、ハードディスクドライブ、CD-ROM、DVD、BLU-RAY、フラッシュドライブ、ディスク、およびその他の既知の物理記憶媒体があげられる。

10

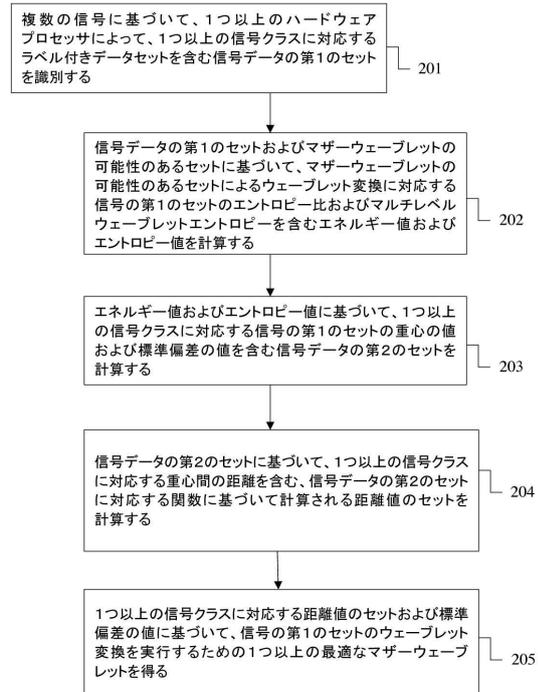
【0041】

開示および例は、例示のみとして考えられ、開示された実施形態の真の範囲および精神は、添付の特許請求の範囲によって示されることが意図される。

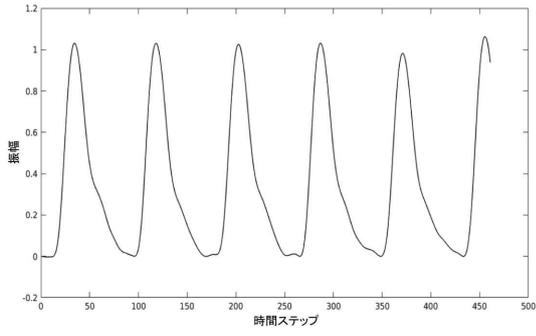
【図1】



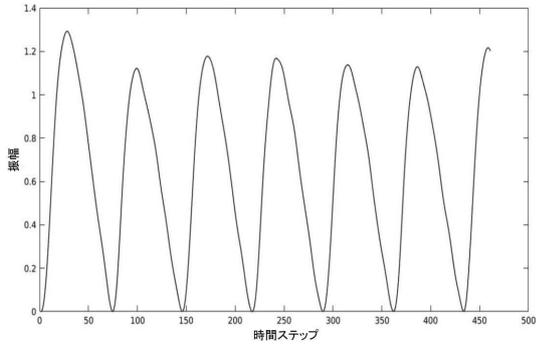
【図2】



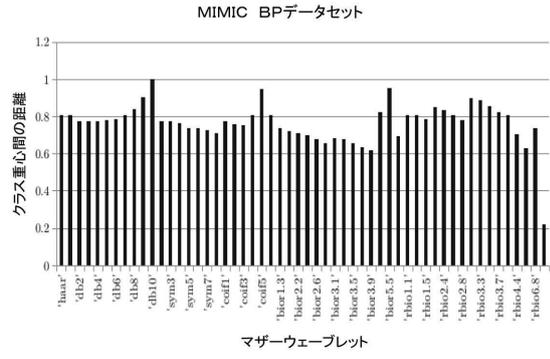
【 図 3 】



【 図 4 】



【 図 5 】



フロントページの続き

- (72)発明者 イシャン サフ
インド国、ウエスト ベンガル、コルカタ、ラジャハット、ニュー タウン、エコスペース プロ
ット - アイアイエフ/12、ビルディング 1ピー、タタ コンサルタンシー サービスズ
リミテッド
- (72)発明者 スネハシス パネルジー
インド国、ウエスト ベンガル、コルカタ、ラジャハット、ニュー タウン、エコスペース プロ
ット - アイアイエフ/12、ビルディング 1ピー、タタ コンサルタンシー サービスズ
リミテッド
- (72)発明者 タヌシュヤム チャットパダイヤイ
インド国、ウエスト ベンガル、コルカタ、ラジャハット、ニュー タウン、エコスペース プロ
ット - アイアイエフ/12、ビルディング 1ピー、タタ コンサルタンシー サービスズ
リミテッド
- (72)発明者 アルパン パル
インド国、ウエスト ベンガル、コルカタ、ラジャハット、ニュー タウン、エコスペース プロ
ット - アイアイエフ/12、ビルディング 1ピー、タタ コンサルタンシー サービスズ
リミテッド
- (72)発明者 ウトパル ガレイン
インド国、ウエスト ベンガル、コルカタ、バラックポア トランク ロード、203、インディ
アン スタティスティカル インスティテュート

審査官 今城 朋彬

- (56)参考文献 BANERJEE, Snehasis et al., Interpretable Feature Recommendation for Signal Analytics ,
arXiv[online], 2017年11月 6日, all 4 pages, URL, <https://arxiv.org/abs/1711.01870>

(58)調査した分野(Int.Cl., DB名)

G 0 6 N 2 0 / 0 0

G 0 6 N 7 / 0 0