

(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号

特許第7122699号
(P7122699)

(45)発行日 令和4年8月22日(2022.8.22)

(24)登録日 令和4年8月12日(2022.8.12)

(51)国際特許分類

F I

G 0 6 Q 50/10 (2012.01)

G 0 6 Q 50/10

G 1 6 C 60/00 (2019.01)

G 1 6 C 60/00

請求項の数 11 (全33頁)

(21)出願番号	特願2018-156142(P2018-156142)	(73)特許権者	314012076
(22)出願日	平成30年8月23日(2018.8.23)		パナソニックIPマネジメント株式会社
(65)公開番号	特開2020-30638(P2020-30638A)		大阪府大阪市中央区城見2丁目1番61号
(43)公開日	令和2年2月27日(2020.2.27)	(74)代理人	100067828
審査請求日	令和3年8月18日(2021.8.18)		弁理士 小谷 悦司
		(74)代理人	100115381
			弁理士 小谷 昌崇
		(74)代理人	100118049
			弁理士 西谷 浩治
		(72)発明者	澤田 好秀
			大阪府門真市大字門真1006番地 パナソニック株式会社内
		(72)発明者	森川 幸治
			大阪府門真市大字門真1006番地 パナソニック株式会社内

最終頁に続く

(54)【発明の名称】 材料情報出力方法、材料情報出力装置、材料情報出力システム、及びプログラム

(57)【特許請求の範囲】

【請求項1】

材料に関する情報を出力する材料情報出力装置における材料情報出力方法であって、
前記材料情報出力装置は、既知組成式リストを学習することで生成された疑似組成式の生成モデルが学習過程で生成した疑似組成式を、学習段階を示す数値と対応付けて予め記憶するメモリを備え、

前記材料情報出力装置のプロセッサが、

前記メモリに記憶された前記疑似組成式の合成可能性を対応する前記数値を用いて算出し、算出した前記合成可能性を前記メモリに記憶された前記疑似組成式の教師ラベルとして付与し、

前記教師ラベルが付与された前記疑似組成式を学習することで前記合成可能性を推定するための推定モデルを生成し、

学習した前記推定モデルを用いて任意の組成式に対する前記合成可能性を算出し、

算出した前記合成可能性を出力する、

材料情報出力方法。

【請求項2】

更に、前記既知組成式リストに含まれる既知組成式のうち推定対象の材料と同じ種類の第1既知組成式を特定し、

予め定められた元素の族の組み合わせに対して、前記組み合わせが一致する前記第1既知組成式を投票し、投票結果を前記メモリに保持し、

前記疑似組成式のうち、前記組み合わせが、前記投票結果が未投票である前記組み合わせと一致する前記疑似組成式に対して、前記合成可能性が0であることを示す前記教師ラベルを付与する、

請求項1に記載の材料情報出力方法。

【請求項3】

前記既知組成式リストは、更に、既知組成式の形成エネルギーを含み、

前記既知組成式リストを用いて学習されたエネルギー推定モデルと、前記既知組成式リストに含まれる形成エネルギーの平均と分散とを前記メモリから取得し、

前記疑似組成式の形成エネルギーを前記エネルギー推定モデルを用いて推定し、前記平均と分散とを用いて前記既知組成式の前記形成エネルギーに対する前記疑似組成式の形成エネルギーの距離を算出し、前記距離が増大するにつれて前記教師ラベルの値が小さくなるように前記教師ラベルを補正する、

請求項1又は2に記載の材料情報出力方法。

【請求項4】

前記メモリは、更に、前記疑似組成式に対する特性値を出力する特性値推定モデルを保持し、

ユーザによって入力された特性値範囲を取得し、

前記特性値範囲内の特性値を前記生成モデルに入力して前記疑似組成式を生成し、

生成した前記疑似組成式の特性値を特性値推定モデルを用いて推定し、

前記合成可能性を示す第1軸と、前記特性値を示す第2軸との少なくとも2軸を備えるグラフ上に、生成した前記疑似組成式について前記合成可能性と前記特性値の推定値とがプロットされた散布図をディスプレイに表示する、

請求項1から3のいずれか1項に記載の材料情報出力方法。

【請求項5】

前記既知組成式リストは、ICSD、materials project、又はNOMADのデータベースで構成される、

請求項1から4のいずれか1項に記載の材料情報出力方法。

【請求項6】

前記メモリは、更に、前記生成モデルが生成した前記疑似組成式が本物であるか偽物であるかを識別する識別モデルを保持し、

前記生成モデル及び前記識別モデルは、敵対的学習アルゴリズムに基づいて学習されたニューラルネットワークで構成されている、

請求項1～5のいずれかに記載の材料情報出力方法。

【請求項7】

前記教師ラベルの付与では、前記学習過程の初期に生成された前記疑似組成式と学習終了時に生成された前記疑似組成式とに対してのみ教師ラベルを付与して前記メモリに保持し、

前記学習段階の初期に生成された前記疑似組成式は、前記合成可能性が0であることを示す数値が付与され、

前記学習終了時に生成された前記疑似組成式は、前記合成可能性が1であることを示す数値が付与される、

請求項1～6のいずれかに記載の材料情報出力方法。

【請求項8】

前記教師ラベルは、前記生成モデルの総学習回数に対する前記学習段階を示す数値の割合が付与される、

請求項1～6のいずれかに記載の材料情報出力方法。

【請求項9】

材料に関する情報を出力する材料情報出力装置であって、

既知組成式リストを学習することで生成された疑似組成式の生成モデルが学習過程で生成した疑似組成式を、学習段階を示す数値と対応付けて予め記憶する第1メモリと、

前記第 1 メモリに記憶された前記疑似組成式の合成可能性を対応する前記数値を用いて算出し、算出した前記合成可能性を前記第 1 メモリに記憶された前記疑似組成式の教師ラベルとして付与する教師ラベル付与部と、

前記教師ラベルが付与された疑似組成式を学習することで前記疑似組成式の合成可能性を推定するための推定モデルを生成する推定モデル学習部と、

前記推定モデルを用いて任意の組成式に対する前記合成可能性を算出する合成可能性算出部と、

前記合成可能性を出力する出力部とを備える材料情報出力装置。

【請求項 10】

請求項 9 記載の材料情報出力装置と、

前記既知組成式を予め記憶する第 2 メモリと、

前記第 2 メモリに記憶された前記既知組成式を学習することで前記生成モデルを生成する生成モデル学習部と、

前記任意の組成式の入力を受け付ける入力部とを備える材料情報出力システム。

【請求項 11】

請求項 9 に記載の材料情報出力装置としてコンピュータを機能させるプログラム。

【発明の詳細な説明】

【技術分野】

【0001】

本開示は、材料に関する情報を出力する技術に関するものである。

【背景技術】

【0002】

近年、人工知能技術を用いて新規材料の候補を材料研究者に提示する研究が盛んに行われている。例えば、特許文献 1 には、問い合わせ蛋白質と任意の化合物とのペアが合成可能か否かを判定する技術が記載されている。

【先行技術文献】

【特許文献】

【0003】

【文献】特許第 5 9 4 6 0 4 5 号公報

【発明の概要】

【発明が解決しようとする課題】

【0004】

しかしながら、特許文献 1 では、実際には合成可能なペアの組成式であるにも拘わらず、合成できないと判定されるケースが起こり得るため、更なる改善が必要である。

【0005】

本開示は、実際には合成可能な未知の組成式が合成不可能と判断されることを防止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】

【0006】

本開示の一態様に係る材料情報出力方法は、材料に関する情報を出力する材料情報出力装置における材料情報出力方法であって、

前記材料情報出力装置は、既知組成式リストを学習することで生成された疑似組成式の生成モデルが学習過程で生成した疑似組成式を、学習段階を示す数値と対応付けて予め記憶するメモリを備え、

前記材料情報出力装置のプロセッサが、

前記メモリに記憶された前記疑似組成式の合成可能性を対応する前記数値を用いて算出し、算出した前記合成可能性を前記メモリに記憶された前記疑似組成式の教師ラベルとして付与し、

前記教師ラベルが付与された前記疑似組成式を学習することで前記合成可能性を推定するための推定モデルを生成し、

10

20

30

40

50

学習した前記推定モデルを用いて任意の組成式に対する前記合成可能性を算出し、算出した前記合成可能性を出力する。

【発明の効果】

【0007】

本開示によれば、実際には合成可能な未知の組成式が合成不可能と判断されることを防止することができる。

【図面の簡単な説明】

【0008】

【図1】実施の形態1に係る材料情報出力システムの構成の一例を示すブロック図である。

【図2】既知組成式リストのデータ構成の一例を示す図である。

【図3】周期表に従って算出される特徴量の一例を示す図である。

【図4】ニューラルネットワークが行う演算の計算モデルを説明するための概念図である。

【図5】ニューラルネットワークが識別問題を学習する様子を説明する図である。

【図6】ニューラルネットワークが回帰問題を学習する様子を説明する図である。

【図7】ニューラルネットワークが生成問題を学習する様子を説明する図である。

【図8】図1に示す生成モデル学習部の詳細な構成の一例を示す図である。

【図9】識別部を構成するニューラルネットワークが学習する様子を説明する図である。

【図10】図1に示す材料情報出力システムの学習過程における処理の一例を示すフローチャートである。

【図11】図1に示す材料情報出力システムの活用過程における処理の一例を示すフローチャートである。

【図12】合成可能性をユーザに提示するための画像の一例を示す図である。

【図13】実施の形態1の変形例に係る材料情報出力システムの構成の一例を示すブロック図である。

【図14】本開示の実施の形態2に係る材料情報出力システムの構成の一例を示すブロック図である。

【図15】図14に示す材料情報出力システムの学習過程における処理の一例を示すフローチャートである。

【図16】投票処理の詳細を示すフローチャートである。

【図17】元素の族を説明する図である。

【図18】実施の形態2に係る既知組成式リストのデータ構成の一例を示す図である。

【図19】本開示の実施の形態3に係る材料情報出力システムの構成の一例を示すブロック図である。

【図20】本開示の実施の形態4に係る材料情報出力システムの構成の一例を示すブロック図である。

【図21】表示部が表示する散布図の一例を示す図である。

【図22】本実施の形態4に係る材料情報出力システムの活用過程における処理の一例を示すフローチャートである。

【図23】材料情報出力システムを実現するためのハードウェア構成の一例を示すブロック図である。

【図24】本開示の材料情報出力システムのハードウェア構成の他の一例を示す図である。

【発明を実施するための形態】

【0009】

（本開示に至る経緯）

近年、材料インフォマティクス分野が脚光を浴びつつある。材料インフォマティクスとは、材料研究の研究スピードを加速させるための、機械学習に代表される人工知能技術（Artificial Intelligence：AI）を指す。材料インフォマティクスでは、機械学習技術によって構築されたモデルを利用して、新規材料の候補を材料研究者に提示し、材料研究者は提示された新規材料の候補の中から所望の材料を作成（合成）する。しかしながら、実際の材料の合成には膨大な時間を必要とするため、提示された

10

20

30

40

50

新規材料の候補の中から所望の材料を実際に作成するのは、手戻りが発生することもあり、非効率的である。そのため、材料インフォマティクスでは、新規材料の候補に対する合成可能性を推定し、新規材料の候補と合わせて材料研究者に提示することが重要となる。

【0010】

上述した特許文献1では、相互作用することが既知の蛋白質と化合物とのペアである第1ペアを正例の教師データとし、蛋白質と化合物とをランダムに組み合わせた第2ペアを負例の教師データとし、サポートベクターマシン等の機械学習技術を用いて、相互作用学習モデルを生成する。そして、この相互作用学習モデルを参照して、問い合わせ蛋白質と化合物とをランダムに組み合わせて得られる複数のペアについての相互作用の可能性を示すスコアを算出する技術が開示されている。

10

【0011】

ところで、蛋白質と化合物とをランダムに組み合わせたペアを組成式として作成した場合、当該組成式は既知ではないが実際には合成可能であるケースも数多くある。

【0012】

しかし、特許文献1では、このような組成式が実際には正例の教師データとして取り扱われるべきところを、負例の教師データとして取り扱われて、相互作用学習モデルが学習されている。そのため、学習後の相互作用学習モデルは、実際には合成可能な未知の組成式の前記スコアを低い値で算出してしまい、この組成式が合成不可能と判定されるという課題がある。

【0013】

本開示は、実際には合成可能な未知の組成式が、合成不可能と判断されることを防止し得る技術を提供する。

20

【0014】

本開示の一態様に係る材料情報出力方法は、材料に関する情報を出力する材料情報出力装置における材料情報出力方法であって、

前記材料情報出力装置は、既知組成式リストを学習することで生成された疑似組成式の生成モデルが学習過程で生成した疑似組成式を、学習段階を示す数値と対応付けて予め記憶するメモリを備え、

前記材料情報出力装置のプロセッサが、

前記メモリに記憶された前記疑似組成式の合成可能性を対応する前記数値を用いて算出し、算出した前記合成可能性を前記メモリに記憶された前記疑似組成式の教師ラベルとして付与し、

30

前記教師ラベルが付与された前記疑似組成式を学習することで前記合成可能性を推定するための推定モデルを生成し、

学習した前記推定モデルを用いて任意の組成式に対する前記合成可能性を算出し、

算出した前記合成可能性を出力する。

【0015】

疑似組成式の生成モデルが学習過程で生成する疑似組成式は学習段階が進むほど合成可能性が高くなる。本構成は、生成モデルが学習過程の後半に生成した疑似組成式ほど高い値の教師ラベルを付与することができ、疑似組成式に対する教師ラベルの振り間違いを抑制できる。そして、このようにして教師ラベルが付与された疑似組成式を学習することで合成可能性を推定するための推定モデルが生成されている。そのため、この推定モデルは、任意の組成式について、実際の合成可能性が高くなるほど、合成可能性の値を高く算出することができる。したがって、本構成は、実際には合成可能な未知の組成式が合成不可能と判断されることを防止できる。

40

【0016】

上記構成において、更に、前記既知組成式リストに含まれる既知組成式のうち推定対象の材料と同じ種類の第1既知組成式を特定し、

予め定められた元素の族の組み合わせに対して、前記組み合わせが一致する前記第1既知組成式を投票し、投票結果を前記メモリに保持し、

50

前記疑似組成式のうち、前記組み合わせが、前記投票結果が未投票である前記組み合わせと一致する前記疑似組成式に対して、前記合成可能性が0であることを示す前記教師ラベルを付与してもよい。

【0017】

本構成によれば、メモリに記憶された疑似組成式のうち、推定対象の材料と同じ種類の第1既知組成式が持つ元素の族の組み合わせとは異なる元素の族の組み合わせを持つ疑似組成式には合成可能性が0であることを示す教師ラベルが付与される。そのため、推定対象の材料とは関連性の低い元素の族の組み合わせを持つ疑似組成式に対して0の教師ラベルを付与して、推定モデルを学習させることができる。その結果、推定対象の材料に関して、推定精度の高い推定モデルを生成できる。

10

【0018】

上記構成において、前記既知組成式リストは、更に、既知組成式の形成エネルギーを含み、

前記既知組成式リストを用いて学習されたエネルギー推定モデルと、前記既知組成式リストに含まれる形成エネルギーの平均と分散とを前記メモリから取得し、

前記疑似組成式の形成エネルギーを前記エネルギー推定モデルを用いて推定し、前記平均と分散とを用いて前記既知組成式の前記形成エネルギーに対する前記疑似組成式の形成エネルギーの距離を算出し、前記距離が増大するにつれて前記教師ラベルの値が小さくなるように前記教師ラベルを補正してもよい。

【0019】

20

組成式は形成エネルギーが高すぎると不安定になり逆に低すぎても不安定となる。本構成によれば、教師ラベルが付与された疑似組成式について、エネルギー推定モデルを用いて形成エネルギーが推定され、推定された形成エネルギーと既知組成式の形成エネルギーとの距離が増大するほど教師ラベルの値が小さく補正される。そして、教師ラベルが補正された疑似組成式を用いて推定モデルが学習されるため、形成エネルギーという物理的特性を考慮に入れて推定モデルを学習させることができ、推定モデルの推定精度を高めることができる。

【0020】

上記構成において、前記メモリは、更に、前記疑似組成式に対する特性値を出力する特性値推定モデルを保持し、

30

ユーザによって入力された特性値範囲を取得し、

前記特性値範囲内の特性値を前記生成モデルに入力して前記疑似組成式を生成し、

生成した前記疑似組成式の特性値を特性値推定モデルを用いて推定し、

前記合成可能性を示す第1軸と、前記特性値を示す第2軸との少なくとも2軸を備えるグラフ上に、生成した前記疑似組成式について前記合成可能性と前記特性値の推定値とがプロットされた散布図をディスプレイに表示してもよい。

【0021】

本構成によれば、ユーザにより特性値範囲が入力されると、その特性値範囲内の特性値を持つ疑似組成式が生成モデルによって生成されると共に、この疑似組成式の特性値が特性値推定モデルによって推定される。そして、この疑似組成式の合成可能性と特性値の推定値とがグラフ上にプロットされた散布図がディスプレイに表示される。そのため、ユーザは、特性値範囲を入力することにより、その特性値範囲内に特性値を持つことが推定される多数の疑似組成式について合成可能性と特性値との関係を容易に把握することができる。

40

【0022】

上記構成において、前記既知組成式リストは、ICSD、materials project、又はNOMADのデータベースで構成されていてもよい。

【0023】

本構成によれば、ICSD、materials project、又はNOMAD等の既存のデータベースが既知組成式リストとして使用されるため、既知組成式リストを作

50

成する手間を省くことができる。

【0024】

上記構成において、前記メモリは、更に、前記生成モデルが生成した前記疑似組成式が本物であるか偽物であるかを識別する識別モデルを保持し、

前記生成モデル及び前記識別モデルは、敵対的学習アルゴリズムに基づいて学習されたニューラルネットワークで構成されていてもよい。

【0025】

本構成によれば、識別モデルと生成モデルとが敵対的学習アルゴリズムに基づいて学習されたニューラルネットワークで構成されているので、生成モデルはより合成可能性の高い疑似組成式を生成することができる。

【0026】

上記構成において、前記教師ラベルの付与では、前記学習過程の初期に生成された前記疑似組成式と学習終了時に生成された前記疑似組成式とに対してのみ教師ラベルを付与して前記メモリに保持し、

前記学習段階の初期に生成された前記疑似組成式は、前記合成可能性が0であることを示す数値が付与され、

前記学習終了時に生成された前記疑似組成式は、前記合成可能性が1であることを示す数値が付与されてもよい。

【0027】

本構成によれば、学習過程の初期に生成された疑似組成式と学習終了時に生成された疑似組成式とに対してのみ教師ラベルが付与されてメモリに保持されるので、メモリの空き容量を確保できる。また、教師ラベルとして0、1の数値のみが使用されているので、データ構造を簡略化し、メモリの空き容量を確保できる。

【0028】

上記構成において、前記教師ラベルは、前記生成モデルの総学習回数に対する前記学習段階を示す数値の割合が付与されてもよい。

【0029】

本構成によれば、教師ラベルを0～1の範囲内の数値で表すことができる。

【0030】

(実施の形態1)

実施の形態1に係る材料情報出力システムは、既知組成式リストに含まれる既知組成式によりニューラルネットワークによって構成される生成モデルを学習させ、学習時に生成モデルから出力される疑似組成式に対し、一定の基準に従って合成可能性を算出し、算出した合成可能性を教師ラベルとして付与する。そして、教師ラベルと疑似組成式とのペアを用いて、ニューラルネットワークによって構成される推定モデルを学習させる。そして、学習された推定モデルに未知の組成式を入力させ、その組成式の合成可能性を出力させる。

【0031】

図1は、実施の形態1に係る材料情報出力システムの構成の一例を示すブロック図である。材料情報出力システムは、材料情報出力装置100と、既知組成式リスト101と、生成モデル学習部102と、入力部103と、記述子算出部108とを備えている。材料情報出力装置100は、疑似組成式保持部104と、教師ラベル付与部105と、推定モデル学習部106と、推定モデル保持部107と、合成可能性算出部109と、表示部110とを備える。材料情報出力システムの各ブロックは、例えば、画像プロセッサ及びマイクロプロセッサ等のプロセッサが所定のプログラムを実行することにより実現されるソフトウェアで構成されてもよい。具体的には、図1において、四角形で示すブロックは、特に断りがなければ、プロセッサで構成され、円筒形で示すブロックは、半導体メモリ等のメモリで構成される。

【0032】

既知組成式リスト101は、ICSD、materials project、及びN

10

20

30

40

50

OMAD等の既知組成式を記憶する既存のデータベース(DB)で構成されている。既知組成式は、材料研究者によってその存在が確認された組成式である。既知組成式リスト101には材料研究者自らが実験を行って存在が確認された組成式が含まれていてもよい。

【0033】

図2は、既知組成式リスト101のデータ構成の一例を示す図である。既知組成式リスト101は、複数の既知組成式のそれぞれについて「組成式」、「特性値」、「特徴量」、及び「記述子」を対応付けて記憶する。「組成式」は、 $\text{LiO}-\text{TiO}_2$ 等の組成式を示す。「特性値」は、電圧及びバッテリー容量等の既知組成式の特性を示すパラメータである。この例では、「特性値」として電圧及びバッテリー容量等の電池に関するパラメータが採用されているが、これは一例であり、電池以外の材料に関するパラメータが採用されてもよい。

10

【0034】

「記述子」は、 $(a, b, c, \alpha, \beta, \gamma)$ で示される材料の結晶構造を表すパラメータで構成される。結晶構造を示すパラメータとしては、例えば、各原子の座標、原子容積、共有結合半径、及び密度等の各元素固有の数値群が採用できる。また、結晶構造を示すパラメータとしては、「 CaMnO_3 」を例に挙げると、Caの原子半径197とMnの原子半径127とOの原子半径60とについて「Ca:Mn:O=1:1:3」の重みを付け平均である「100.8」という値が採用されてもよい。

【0035】

「特徴量」は、例えば、図3に示すように周期表における元素の割合に基づいて算出される数値群(ベクトル)によって表現される。図3は、周期表に従って算出される特徴量の一例を示す図である。以下、 $\text{LiCo}(\text{PO}_3)_4$ で表される組成式301を例に挙げて特徴量304の算出について説明する。まず、この組成式301は、周期表302の元素を左上から右下へラスターキャンすることで次元のベクトル303に変換される。ベクトル303は、周期表の順番で元素が対応付けられた複数のベクトル要素で構成されている。そして、組成式301を構成する元素の元素数が、対応するベクトル要素に割り当てることでベクトル303が作成される。 $\text{LiCo}(\text{PO}_3)_4$ はLi、Co、O、Pの元素からなり、それぞれの元素数は、「1」、「1」、「12」、「4」であるため、ベクトル303は、 $(0, 0, 1, \dots, 12, \dots, 4, \dots, 1, 0 \dots)$ で表される。次に、ベクトル303は、ベクトル要素の総和が1となるように正規化され、特徴量304が算出される。

20

30

【0036】

なお、これら特徴量及び記述子は、実際の組成式への変換が容易な値となっているため、特徴量及び記述子は、組成式と一対一対応する。したがって、特徴量及び記述子の少なくとも一方が分かれば組成式を一意的に特定できる。そこで、以下の説明において、疑似組成式を生成することは、特徴量及び記述子の少なくとも一方を生成することと同じであるものとする。

【0037】

図1に参照を戻す。生成モデル学習部102は、上記の既知組成式リスト101に含まれる既知組成式を用いて組成式を生成するための生成モデルを学習する。生成モデルから出力される組成式には、既知組成式リスト101には含まれない組成式も含まれる。本実施の形態では、学習時に付随的に生成される組成式を疑似組成式と呼び、それを用いて材料の合成可能性を推定するための推定モデルを構築する。

40

【0038】

以下、生成モデル学習部102を構成するニューラルネットワークについて説明する。図4は、ニューラルネットワークが行う演算の計算モデルを説明するための概念図である。ニューラルネットワークは、周知のように、生物のニューラルネットワークを模した計算モデルに従って演算を行う演算装置である。

【0039】

図4に示されるように、ニューラルネットワーク400は、ニューロンに相当する複数

50

のユニット（白丸で示されている）を、入力層 4 0 1、隠れ層 4 0 2、及び出力層 4 0 3 のそれぞれに配置して構成される。

【0 0 4 0】

隠れ層 4 0 2 は、図 4 に示す例では、2 つの隠れ層 4 0 2 a 及び隠れ層 4 0 2 b で構成されているが、単一の隠れ層若しくは 3 つ以上の隠れ層で構成されてもよい。なお、複数の隠れ層を有するニューラルネットワークは、特に、多層ニューラルネットワーク装置と呼ばれることがある。

【0 0 4 1】

ここで、入力層 4 0 1 に近い層を下位層とし、出力層 4 0 3 に近い層を上位層とする。この場合、各ユニットは、下位層に配置されたユニットから受信した計算結果を荷重値に応じて結合（例えば、荷重和演算）し、該当結合の結果を上位層に配置されたユニットに送信する計算要素である。

10

【0 0 4 2】

ニューラルネットワーク 4 0 0 の機能は、ニューラルネットワーク 4 0 0 が有する層の数及び各層に配置されるユニットの数を表す構成情報と、各ユニットでの荷重和演算に用いられる荷重値を表す荷重 $W = [w_1, w_2, \dots]$ とで定義される。

【0 0 4 3】

図 4 に示すように、ニューラルネットワーク 4 0 0 は、入力層 4 0 1 の各入力ユニット 4 0 5 に入力データ $X = [x_1, x_2, \dots]$ （画像であれば画素値）が入力される。これにより、隠れ層 4 0 2 の隠れユニット 4 0 6 及び出力層 4 0 3 の出力ユニット 4 0 7 において荷重 $W = [w_1, w_2, \dots]$ を用いた荷重和演算が行われ、出力層 4 0 3 の各出力ユニット 4 0 7 の出力値から出力データ $Y = [y_1, y_2, \dots]$ が出力される。

20

【0 0 4 4】

入力データ X がニューラルネットワーク 4 0 0 に入力された際に所望の出力データ Y を出力するように、ニューラルネットワーク 4 0 0 の荷重 W の値を調整することをニューラルネットワークの学習と呼ぶ。ニューラルネットワーク 4 0 0 を学習する場合、例えば、入力データ X 及び出力データ Y との誤差を表す損失関数を定義し、勾配降下法により当該損失関数を減少させる勾配に沿って荷重 W は更新される。

【0 0 4 5】

ニューラルネットワークを含む機械学習の問題は、出力データの種類に応じて識別問題、回帰問題、及び生成問題の 3 つの問題に分けられる。まず、識別問題について説明する。図 5 は、ニューラルネットワークが識別問題を学習する様子を説明する図である。識別問題は与えられた画像が「犬」か「猫」かを判定するようにニューラルネットワークの荷重 W の値を求める問題のことを指す。

30

【0 0 4 6】

より具体的には、ニューラルネットワーク 4 0 0 では、各出力ユニットは入力データ X を分類するための異なる正解ラベルに対応付けられている。また、荷重 W は複数の入力データ X の各々が入力されたときに、当該入力データ X の正しい正解ラベルに対応する出力ユニットの出力値が 1 に近づき、他の出力ユニットの出力値が 0 に近づくように調整される。

40

【0 0 4 7】

つまり、図 5 に示される例では、ニューラルネットワーク 4 0 0 において、各出力ユニットは、例えば、「犬」であれば正解ラベルが「1、0」、「猫」であれば正解ラベルが「0、1」というように、「犬」及び「猫」を示す正解ラベルのうち異なる 1 つの正解ラベルが対応付けられる。また、荷重 W は、「犬」の正解ラベルが付与された学習データが入力されると出力データ [1、0] を出力し、「猫」が付与された学習データが入力されると出力データ [0、1] を出力するように調整される。

【0 0 4 8】

なお、ニューラルネットワーク 4 0 0 が多層ニューラルネットワークである場合、特に、前記教師付き学習を行う前に、「layer-wise pre-training」

50

と呼ばれる学習手法によって、荷重値は隠れ層ごとに個別に調整されても良い。これにより、その後の学習において、汎化性能が高い荷重 W が得られる。

【0049】

また、ニューラルネットワーク400の荷重値の調整には、上述した勾配降下法の他にも、例えば、「Adam」などの周知のアルゴリズムが用いらてもよい。また、ニューラルネットワーク400が多層ニューラルネットワークである場合、各層で個別に学習が実行されても良い。

【0050】

次に、回帰問題について説明する。図6は、ニューラルネットワークが回帰問題を学習する様子を説明する図である。回帰問題は「犬」及び「猫」といった正解ラベルとは異なり、例えば、材料の「特性値」のような実数値が推定されるように、ニューラルネットワークの荷重 W の値を求める問題のことを指す。

10

【0051】

より具体的には、ニューラルネットワーク400では、出力ユニットのそれぞれには、出力データ Y を構成する実数値 y_1 、 y_2 が対応付けられ、入力ユニットのそれぞれには、入力データ X を構成する実数値 x_1 、 x_2 が対応付けられている。荷重 W は、複数の入力データ X の各々が入力されたときに、当該入力データ X に対応する正しい実数値が各出力ユニットから出力されるように調整される。

【0052】

図6の例では、ニューラルネットワーク400において、出力データ Y は、実数値 y_1 が例えば電圧、実数値 y_2 が例えば容量というように、材料の特性値で構成されている。また、入力データ X は、 x_1 が例えばパラメータ α 、 x_2 がパラメータ β 、 x_3 が特徴量というように、材料の「特徴量及び記述子」で構成されている。ここでは、「特徴量及び記述子」と記載したが、これらは少なくとも一方が用いられてもよい。記述子は、図2で説明した材料の結晶構造を示すパラメータである。特徴量は、図3に示すように元素の割合によって算出される一次元ベクトルである。なお、荷重 W は、ニューラルネットワーク400から出力（推定）された特性値が正解特性値に近づくように調整される。

20

【0053】

次に、生成問題について説明する。図7は、ニューラルネットワークが生成問題を学習する様子を説明する図である。生成問題は、上記の識別問題又は回帰問題とは異なり、既知DBに含まれない何らかのデータを生成する問題である。例えば、ある特性値を持つ組成式を複数生成するというように、生成問題は回帰問題の逆問題に対応する。加えて、識別問題及び回帰問題は入力と出力とが1対1対応となっている必要があるが、生成問題は1つの入力データ X に対して複数の出力データ Y が出力されてもよい。これは、ある特性値を満たす組成式は複数存在しうるからである。図1に示す生成モデル学習部102は、この生成問題を解くニューラルネットワークを学習する。

30

【0054】

より具体的には、図7に示すニューラルネットワーク400では、出力ユニットのそれぞれには出力データ Y を構成する実数値 y_1 、 y_2 が対応付けられ、入力ユニットのそれぞれには入力データ X を構成する実数値 x_1 、 x_2 、・・・対応付けられている。また、荷重 W は、複数の入力データ X の各々が入力されたときに、複数の入力データ X のそれぞれに対応する正しい実数値が出力されるように調整される。例えば、組成式「LiO₂」の特性値である電圧「9.43」及び容量「1214」が入力データ X [x_1 、 x_2]として入力されると、組成式「LiO₂」の「特徴量及び記述子」が出力されるように荷重 W が調整される。

40

【0055】

このとき、入力データ X は、識別問題又は回帰問題における出力データ Y に対応し、出力データ Y は識別問題又は回帰問題における入力データ X に対応する。

【0056】

つまり、図7に示すニューラルネットワーク400において、出力データ Y は、 y_1 が

50

例えば組成式「 $\text{LiO}-1\text{TiO}_2$ 」の記述子のパラメータ α 、 y_2 が例えば組成式「 $\text{LiO}-1\text{TiO}_2$ 」の記述子のパラメータ β というように、材料の特微量及び記述子で構成されている。また、入力データ X は、 x_1 が例えば組成式「 $\text{LiO}-1\text{TiO}_2$ 」の電圧、 x_2 が例えば組成式「 $\text{LiO}-1\text{TiO}_2$ 」の容量というように、材料の特性値で構成されている。

【0057】

そして、荷重 W は、ニューラルネットワーク400から出力（推定）された「特微量及び記述子」が正解に近づくように調整される。

【0058】

前述のように、生成モデル学習部102において、ニューラルネットワーク400は、
入力データ X から出力データ Y を生成する生成問題を解く。この生成問題は、例えば、敵対的学習アルゴリズムによって導出される。敵対的学習アルゴリズムは、「generative adversarial network (GAN)」とも称され、近年盛んに研究が行われている。

【0059】

以下、組成式の生成を題材として、材料応用へ適している、GANの一種であるconditional-GAN (C-GAN) について説明する。C-GANは、生成部と識別部とのそれぞれを構成する2種類のニューラルネットワークを利用して生成部のニューラルネットワークを学習させるアルゴリズムである。図8は、図1に示す生成モデル学習部102の詳細な構成の一例を示す図である。

【0060】

生成モデル学習部102は、乱数発生部801、生成部802、及び識別部803を備える。乱数発生部801は、所望数の乱数を発生させる。この乱数の発生数が生成部802にて生成される疑似組成式の数となる。例えば、乱数の発生数が1000個の場合、1000個の疑似組成式が生成される。

【0061】

次に、生成部802は、各乱数と既知組成式リスト101に含まれる既知組成式の各特性値とのペアをニューラルネットワークNN1に入力する。すなわち、ニューラルネットワークNN1の入力データ X は $X = [\text{乱数}, \text{特性値}]$ となる。この入力データ X が与えられたとき、ニューラルネットワークNN1は荷重 W の値に従って出力データ Y を算出する。この出力データ Y が疑似組成式に対応する。ここで、疑似組成式とは、図2に示す記述子、図3に示す特微量、又はこれらの組み合わせに対応し、元の組成式への変換が容易な量を指す。

【0062】

識別部803は、生成部802で生成された疑似組成式の「特微量及び記述子」が、実際の組成式の「特微量及び記述子」と近い値を生成できているかを判定する。この判定を行うために、識別部803を構成するニューラルネットワークNN2は、図9に示すように、分類問題及び回帰問題の両問題を解く。

【0063】

図9は、識別部803を構成するニューラルネットワークNN2が学習する様子を説明する図である。図9において、ニューラルネットワーク400は、ニューラルネットワークNN2に対応する。ニューラルネットワーク400には、生成部802が生成した疑似組成式の「特微量及び記述子」が入力データ X として入力される。また、ニューラルネットワーク400には、既知組成式リスト101に含まれる既知組成式の「特微量及び記述子」が入力データ X として入力される。

【0064】

分類問題においては、ニューラルネットワーク400は、既知組成式リスト101に含まれる既知組成式の「特微量及び記述子」を「本物」と判別し、疑似組成式の「特微量及び記述子」を「偽物」と判別する二値判別を実施する。すなわち、ニューラルネットワーク400は、本物ラベルを持つ「特微量及び記述子」が入力された場合、出力データ $Y =$

10

20

30

40

50

[1, 0] を出力し、偽物ラベルを持つ「特徴量及び記述子」が入力された場合、出力データ $Y = [0, 1]$ を出力するように荷重 W の値を調整する。ここで、[1, 0] は本物ラベルに対応する出力データ Y を表し、[0, 1] は偽物ラベルに対応する出力データ Y を表す。

【0065】

一方、回帰問題においては、ニューラルネットワーク 400 は、既知組成式の「特徴量及び記述子」が入力された場合、特性値を正しく推定できるように荷重 W の値を調整する。すなわち、識別部 803 を構成するニューラルネットワーク NN2 は、「二値判別＋特性値推定」の両方が正しく行えるように荷重 W を調整する。

【0066】

図 8 に参照を戻す。識別部 803 での出力データ (Y [本物、偽物]、 Y [特性値]) は、生成部 802 へ入力され、生成部 802 を構成するニューラルネットワーク NN1 の荷重 W の調整にも利用される。生成部 802 は、識別部 803 において疑似組成式が「本物」と間違われるようにニューラルネットワーク NN1 の荷重 W を調整する。

【0067】

出力データ (Y [本物、偽物]、 Y [特性値]) が入力されると、生成部 802 は、出力データ Y [本物、偽物] と本物ラベルに対応する出力データ Y [1, 0] との距離 D を算出する。例えば、出力データ Y [本物、偽物] として $Y = [0.8, 0.2]$ が入力された場合、生成部 802 は、距離 D として $D = \sqrt{(1 - 0.8) \times (1 - 0.8) + (0 - 0.2) \times (0 - 0.2)} = 0.3$ を算出する。一方、出力データ Y [本物、偽物] として、 $Y = [0.2, 0.8]$ が入力された場合、生成部 802 は、距離 D として $D = \sqrt{(1 - 0.2) \times (1 - 0.2) + (0 - 0.8) \times (0 - 0.8)} = 1.1$ を算出する。この場合、出力データ Y [0.8, 0.2] の方が出力データ Y [0.2, 0.8] よりも距離 D が小さいため、本物に近いことが分かる。そのため、出力データ Y [0.8, 0.2] に対応する疑似組成式の方が、出力データ Y [0.2, 0.8] に対応する疑似組成式よりも、生成部 802 の生成結果として好ましい。そこで、生成部 802 は、距離 D が小さくなるように、ニューラルネットワーク NN1 の荷重 W を調整する。

【0068】

加えて、生成部 802 は、識別部 803 からの出力データ Y [特性値] も入力されている。この出力データ Y [特性値] は、生成部 802 が生成した疑似組成式に対する識別部 803 の特性値の推定結果を表している。識別部 803 からの出力データ Y [特性値] が、対応する疑似組成式の生成に用いられた特性値 (既知組成式の特性値) と近い値になっているか否かも重要な情報である。すなわち、ある疑似組成式に対する識別部 803 の特性値の推定結果がその疑似組成式の生成に用いられた特性値に近いほど、生成部 802 はより本物に近い (実在する可能性の高い) 疑似組成式を生成できていると言える。

【0069】

そこで、生成部 802 は、識別部 803 からの出力データ Y [特性値] と、出力データ Y [特性値] に対応する疑似組成式を生成する際に用いた特性値との差 (距離) が小さくなるように、ニューラルネットワーク NN1 の荷重 W を調整する。このように、生成部 802 は、上記 2 種類の距離を用いてニューラルネットワーク NN1 の荷重 W を調整する。

【0070】

これによって、生成部 802 を構成するニューラルネットワーク NN1 は、入力された特性値を持つような疑似組成式を生成することができるようになる。

【0071】

このとき、識別部 803 を構成するニューラルネットワーク NN2 は、二値判別と特性値推定の 2 つの問題 (タスク) を同時に解くことになる。これは、マルチタスク学習と呼ばれる技術である。マルチタスク学習は、同時に異なるタスクを解くことで、タスク間の知見を共有することができ、各タスクの精度を向上させることができる。

【0072】

10

20

30

40

50

生成部 802 が生成する疑似組成式の精度は学習回数が増大するにつれて徐々に向上していくため、識別部 803 が解く問題の難易度は徐々に上昇していく。簡単な問題からニューラルネットワークに解かせ、徐々に問題の難易度を上昇させていく学習方法はカリキュラム学習と呼ばれている。したがって、本実施の形態では、識別部 803 にカリキュラム学習を実施させてもよい。カリキュラム学習では、簡易な問題が難易度の高い問題に対する制約の役割を果たす。そのため、識別部 803 は判別精度及び識別精度の高いニューラルネットワーク NN2 を生成できる。

【0073】

このように、マルチタスク学習及びカリキュラム学習を実施することによって、識別部 803 は既知 DB 外のデータに対する汎化性能が向上する。そして、識別部 803 の汎化性能が高まると、生成部 802 の汎化性能も高めることができる。その結果、生成部 802 は、既知 DB 内に含まれない特性値が入力されたとしても、識別部 803 のニューラルネットワーク NN2 が「本物」と誤判定してしまうような疑似組成式が生成できる。

【0074】

図 1 に参照を戻し、入力部 103 について説明する。入力部 103 は、乱数発生部 801 における乱数の発生数、すなわち、生成したい疑似組成式の数指定するユーザからの入力を受け付ける。また、入力部 103 は、合成可能性の評価対象となる組成式を指定するユーザからの入力を受け付ける。入力部 103 は、例えば、キーボード、及びマウスといった入力装置で構成される。

【0075】

次に、材料情報出力装置 100 について説明する。疑似組成式保持部 104 は、生成モデル学習部 102 が生成した疑似組成式を保持する。疑似組成式は、生成モデル学習部 102 が備える生成部 802 のニューラルネットワーク NN1 の出力データ Y のことを指し、「特徴量及び記述子」で構成される。通常、疑似組成式保持部 104 は、学習完了後に生成部 802 により生成された疑似組成式を保持するが、本実施の形態においては、学習過程において生成部 802 により生成された疑似組成式も保持する。

【0076】

学習過程においては、まず、生成部 802 及び識別部 803 のニューラルネットワーク NN1, NN2 の荷重 W は、初期値が設定される。その後、初期値が設定されたニューラルネットワーク NN1, NN2 は、前述した敵対的学習アルゴリズムにより学習を進めていく。このとき、生成部 802 は、ニューラルネットワーク NN1 が各学習段階で生成した疑似組成式を学習段階を示す数値と対応付けて疑似組成式保持部 104 に保持させる。なお、ニューラルネットワーク NN1, NN2 の荷重 W は、例えば、勾配降下法等で調整される。

【0077】

例えば、乱数の発生数が 1000 であるとする、ニューラルネットワーク NN1 は、1 回の学習を行う度に 1000 個の疑似組成式を生成する。したがって、総学習回数が T (T は 2 以上の整数) 回であるとする、生成部 802 は、0 回目の学習で生成した 1000 個の疑似組成式、1 回目の学習で生成した 1000 個の疑似組成式、・・・、T 回目の学習で生成した 1000 個の疑似組成式というように、 $1000 \times (T + 1)$ 個の疑似組成式を学習段階を示す数値 (0 回目、1 回目、・・・、T 回目) と対応付けて疑似組成式保持部 104 に保持させる。なお、0 回目の学習においては、荷重 W として初期値が設定されたニューラルネットワーク NN1 に対して、乱数及び既知組成式の特徴値の 1000 個のペアが入力され、各ペアに対応する 1000 個の疑似組成式が生成される。

【0078】

教師ラベル付与部 105 は、疑似組成式保持部 104 に保持されている各疑似組成式の合成可能性を、各疑似組成式に対応付けられた学習段階を示す数値を用いて算出し、算出した合成可能性を疑似組成式保持部 104 が保持する疑似組成式の教師ラベルとして付与する。

【0079】

ここで、教師ラベル付与部 105 は、総学習回数 T に対する学習段階を示す数値 t の割合を教師ラベルとして付与する。この場合、教師ラベルは $0 \sim 1$ の値を持つ 1 次元の実数値となる。例えば 3 回目の学習で得られた疑似組成式の教師ラベル L は $L = [3 / T]$ となる。しかし、この教師ラベルの付与方法では、異なる数値 t に対して同一の疑似組成式が生成された場合、同一の疑似組成式に対して複数の教師ラベル L が付与されてしまう。

【0080】

そこで、本実施の形態では、同一の疑似組成式に対して複数の教師ラベル L が付与された場合、教師ラベル付与部 105 は、当該疑似組成式に対して、最後に生成されたときの教師ラベル L 、すなわち、最大の教師ラベル L を付与する。例えば、ある疑似組成式が、5 回目、20 回目、80 回目のそれぞれの学習段階で生成されたとすると、その疑似組成式に対して付与される教師ラベル L は $L = 80 / T$ となる。これにより、教師ラベル付与部 105 は、疑似組成式と教師ラベル L とを一対一で対応付けることができる。

【0081】

推定モデル学習部 106 は、教師ラベル付与部 105 により教師ラベルが付与された疑似組成式を学習することで、任意の組成式の合成可能性を推定するための推定モデルを生成する。ここで、推定モデル学習部 106 は、疑似組成式が何回目の学習で得られた疑似組成式であるかを示す $0 \sim 1$ の実数値を推定する回帰問題を解くことによって推定モデルを生成する。すなわち、推定モデル学習部 106 は、入力データ X が疑似組成式の記述子及び特徴量、出力データ Y が教師ラベル L を示す実数値となる。

【0082】

以下、疑似組成式が何回目の学習段階で得られたものであるかを推定することが合成可能性を判定することに一致する理由について説明する。

【0083】

前述した C-GAN は、既知組成式リスト 101 を用いて学習される。既知組成式リスト 101 は、既に存在が確認された既知組成式のリストである。すなわち、合成可能性の範囲を 0 （合成可能性低） ~ 1 （合成可能性高）の値とした場合、既知組成式リスト 101 に含まれる既知組成式の合成可能性は 1 となる。一方、生成部 802 を構成するニューラルネットワーク NN1 は、学習の初期段階では既知組成式と大きく異なる疑似組成式を生成する。しかし、徐々に学習を繰り返していくことにより、ニューラルネットワーク NN1 は、既知組成式に似通った疑似組成式を生成する。

【0084】

このことから、学習の初期段階に得られる疑似組成式は合成可能性が低い、学習の終了間際に得られる疑似組成式は合成可能性が高いと言える。従って、初期の学習で得られた疑似組成式、すなわち、教師ラベル L として、 $L = [0 / T] = [0]$ が付与された疑似組成式は、最も合成可能性が低いことを表すのである。一方、 T 回目の学習で得られた疑似組成式、すなわち、教師ラベル L として、 $L = [T / T] = [1]$ が付与された疑似組成式は最も合成可能性が高いことを表すのである。

【0085】

上述した教師ラベル L の振り方は、合成可能性と学習段階を示す数値 t との関係が線形であると仮定している。但し、これは一例であり、本実施の形態においては、合成可能性と数値 t との関係は線形でなく、非線形であってもよい。例えば、 t 回目の学習で得られた疑似組成式の合成可能性 Z とすると、例えば、 $Z = \text{sigmoid}(t)$ により合成可能性は決定されてもよい。ここで、 $\text{sigmoid}(t)$ は、シグモイド関数を表す。この場合、教師ラベル付与部 105 は、合成可能性 Z を疑似組成式に対する教師ラベル L として付与すればよい。

【0086】

本実施の形態では、推定モデル学習部 106 は、正例負例の二値判別ではなく、回帰問題として合成可能性を推定する。加えて、教師ラベル L はニューラルネットワーク NN1 の学習段階を示す数値 t に基づいて付与される。生成問題を解くニューラルネットワーク NN1 は、学習を繰り返すことによって、識別部 803 を構成するニューラルネットワー

10

20

30

40

50

クNN2が本物の組成式と見分けがつかないような疑似組成式を生成していく。すなわち、生成部802が生成する疑似組成式は、合成可能性が段階的に高まっていく。

【0087】

さらに、本実施の形態においては、学習過程において、同一の疑似組成式が複数生成された場合、当該疑似組成式に対して最後に生成されたときの教師ラベルLが付与される。例えば、学習過程の初期段階で合成可能性が高い疑似組成式が生成されたとする。この疑似組成式は、合成可能性が高いため、学習の中期～後期、終了間際においても高確率で生成部802により生成される。加えて、乱数の発生数を多くすればするほど、この疑似組成式の生成部802による発生確率は高まる。以上のことから、疑似組成式に対して最大の教師ラベルLを付与することで、学習過程における合成可能性を示す教師ラベルの振り間違いによる学習効率の低下を抑制することができる。

10

【0088】

さらに、本実施の形態においては、推定モデル学習部106は、推定モデルの学習に生成部802にて生成された疑似組成式を利用している。生成部802は設定された乱数の発生数だけ疑似組成式を生成する。そのため、推定モデルの学習に必要なデータを大量に準備することができる。加えて、生成部802にて生成される疑似組成式は、既知組成式リスト101に含まれていない、すなわち、既知DBに存在しない疑似組成式を多数含んでいる。これらにより、推定モデル学習部106は、推定モデルの過学習を抑制することができる。

【0089】

推定モデル保持部107は、推定モデル学習部106が生成した推定モデルを保持する。推定モデル保持部107により保持される情報としては、推定モデルを構成するニューラルネットワークの層の数及び各層に配置されたユニットの数を表す構成情報と、各ユニットでの荷重和計算に用いられる荷重値を表す荷重Wとが含まれる。

20

【0090】

活用過程では、入力部103には合成可能性の評価対象となる組成式がユーザにより入力される。記述子算出部108は、入力部103に入力された組成式を特徴量及び記述子に変換し、合成可能性算出部109に入力する。ここで、記述子算出部108は、組成式の結晶構造を表すパラメータを算出するための所定の数式に評価対象となる組成式を入力することで記述子を算出すればよい。また、記述子算出部108は、評価対象となる組成式について、図3で示す一次元のベクトル303を算出することで、当該組成式の特徴量を算出ればよい。

30

【0091】

合成可能性算出部109は、推定モデル保持部107に保持されている構成情報と荷重Wとで表される推定モデルを読み出す。そして、合成可能性算出部109は、推定モデルに対して、記述子算出部108により算出された特徴量及び記述子を入力データXとして入力する。そして、合成可能性算出部109は、入力データXが入力ユニットに与えられたときの推定モデルを構成する各ユニットでの荷重和を算出し、合成可能性を表す出力データYを出力する。

【0092】

表示部110は、例えば、液晶表示ディスプレイ等の表示装置及びプロセッサで構成され、合成可能性算出部109によって算出された合成可能性を示す画像を生成して表示する。

40

【0093】

図10は、図1に示す材料情報出力システムの学習過程における処理の一例を示すフローチャートである。

【0094】

S101では、生成モデル学習部102は、既知組成式リスト101に含まれる既知組成式を読み込む。S102では、生成モデル学習部102は、乱数発生部801で生成された乱数とS101で読み込んだ既知組成式の特徴値とのペアを生成部802のニューラ

50

ルネットワーク NN 1 に入力し、疑似組成式を得る。例えば、乱数の発生数が 1 0 0 0 であるとする、1 0 0 0 個のペアがニューラルネットワーク NN 1 に入力されて 1 0 0 0 個の疑似組成式が生成される。

【0095】

S 1 0 3 では、生成モデル学習部 1 0 2 の生成部 8 0 2 は、ニューラルネットワーク NN 1 によって生成された疑似組成式を学習段階を示す数値 t と対応付けて疑似組成式保持部 1 0 4 に保持させる。例えば、上記の 1 0 0 0 個の例では、1 回の学習につき 1 0 0 0 個の疑似組成式が生成されるので、1 つの数値 t に対して 1 0 0 0 個の疑似組成式が対応付けて保持される。

【0096】

S 1 0 4 では、生成モデル学習部 1 0 2 の識別部 8 0 3 は、疑似組成式と既知組成式とをニューラルネットワーク NN 2 に入力する。例えば、識別部 8 0 3 は、 t 回目の学習において生成部 8 0 2 で生成された 1 0 0 0 個の疑似組成式と、既知組成式リスト 1 0 1 に含まれる既知組成式とを交互にニューラルネットワーク NN 2 に入力すればよい。

【0097】

その後、S 1 0 5 から S 1 0 7 の処理が並列で実施される。

【0098】

S 1 0 5 では、識別部 8 0 3 は、ニューラルネットワーク NN 2 に入力された既知組成式の実験値を推定し、実験の実験値との誤差を計算する。実験の実験値とは、実験値が推定された既知組成式に対応する既知組成式リスト 1 0 1 が保持する実験値のことを指す。

【0099】

S 1 0 6 では、識別部 8 0 3 は、ニューラルネットワーク NN 2 に入力された既知組成式と疑似組成式とのそれぞれについて二値判別を実施し、正答率を算出する。

【0100】

S 1 0 7 では、識別部 8 0 3 は、ニューラルネットワーク NN 2 に入力された疑似組成式の実験値を推定し、推定した実験値と、S 1 0 2 において乱数とペアで入力された実験値との誤差を算出する。ここで、乱数とペアで入力された実験値とは、実験値が推定された疑似組成式を生成する際に生成部 8 0 2 に入力された既知組成式の実験値のことを指す。

【0101】

S 1 0 8 では、識別部 8 0 3 は、S 1 0 5 で算出された誤差が小さくなり、且つ、S 1 0 6 で算出された正答率が高くなるように、ニューラルネットワーク NN 2 の荷重 W を更新する。

【0102】

S 1 0 9 では、生成部 8 0 2 は、S 1 0 6 で算出された正答率が低くなり、且つ、S 1 0 7 で算出された誤差が小さくなるように、ニューラルネットワーク NN 1 の荷重 W を更新する。ここで、識別部 8 0 3 は、疑似組成式の二値判別結果は生成部 8 0 2 に出力するが、既知組成式の二値判別結果は生成部 8 0 2 に出力しない。したがって、生成部 8 0 2 には、識別部 8 0 3 の疑似組成式に対する二値判別結果を示す出力データ Y [本物、偽物] が入力される。

【0103】

生成部 8 0 2 が生成した疑似組成式に対して、識別部 8 0 3 が本物と判別した場合、すなわち、識別部 8 0 3 が誤判別をした場合、生成部 8 0 2 は、より本物に近い疑似組成式を生成できていると言える。したがって、疑似組成式に対する出力データ Y [本物、偽物] と正解ラベル [1, 0] との距離 D が小さいほど、生成部 8 0 2 は、より本物に近い疑似組成式を生成できていると言える。そこで、生成部 8 0 2 は、距離 D が小さくなるようにニューラルネットワーク NN 1 の荷重 W を調整することで、S 1 0 6 で算出された正答率を低くしている。

【0104】

S 1 1 0 では、生成モデル学習部 1 0 2 は、ニューラルネットワーク NN 1, NN 2 の学習回数が総学習回数 T に到達すると、処理を S 1 1 1 に遷移させ、総学習回数 T に到達

10

20

30

40

50

していなければ、処理を S 1 0 2 に遷移させる。

【0105】

S 1 1 1 では、教師ラベル付与部 1 0 5 は、疑似組成式保持部 1 0 4 に保持されている疑似組成式に対して、教師ラベル L を付与する。

【0106】

S 1 1 2 では、推定モデル学習部 1 0 6 は、疑似組成式から教師ラベル L の値を推定する回帰問題を解くニューラルネットワークを学習する。S 1 1 3 では、推定モデル学習部 1 0 6 は、S 1 1 2 で学習されたニューラルネットワークを推定モデルとして推定モデル保持部 1 0 7 に保持させる。以上により学習過程が終了される。

【0107】

図 1 1 は、図 1 に示す材料情報出力システムの活用過程における処理の一例を示すフローチャートである。S 2 0 1 では、入力部 1 0 3 は、合成可能性の評価対象となる組成式が入力される。S 2 0 2 では、合成可能性算出部 1 0 9 は、推定モデルを推定モデル保持部 1 0 7 から読み込む。S 2 0 3 では、記述子算出部 1 0 8 は、S 2 0 1 にて入力された組成式の特徴量及び記述子を算出する。

【0108】

S 2 0 4 では、合成可能性算出部 1 0 9 は、S 2 0 3 で算出された特徴量及び記述子を推定モデルに入力し、合成可能性を算出する。S 2 0 5 では、表示部 1 1 0 は、算出された合成可能性を示す画像を表示する。以上により活用過程が終了される。

【0109】

図 1 2 は、合成可能性をユーザに提示するための画像 G 1 の一例を示す図である。画像 G 1 には、ユーザが入力した組成式に対する合成可能性の数値が表示されている。これにより、ユーザは、評価対象となる組成式の合成可能性を容易に認識できる。

【0110】

以上説明したように、実施の形態 1 に係る材料情報出力システムによれば、疑似組成式に対して適切な合成可能性を示す教師ラベル L を付与することができるので、教師ラベル L の振り間違いを抑制することができる。そして、適切な教師ラベル L が付与された疑似組成式を用いて学習された推定モデルを用いて合成可能性が算出されているため、合成可能性の適切な値を材料研究者に提示することができる。その結果、実際には合成可能な未知の組成式が合成不可能と判断されることを防止できる。

【0111】

また、本実施の形態における材料情報出力装置は、推定モデルの構築に既知組成式リスト 1 0 1 を明示的に利用していない。すなわち、教師ラベル L の付与されていない、未知の組成式を含む疑似組成式が入力として与えられ、合成可能性を推定する推定モデルが学習される。従って、推定モデルを学習する際には、既知組成式リスト 1 0 1 を保持しておく必要はない。これにより、既知組成式リスト 1 0 1 を保持する企業は、企業内の実験データ等の秘匿情報として既知組成式リスト 1 0 1 を社内に残しつつ、推定モデルの学習を行う材料情報出力装置 1 0 0 は外部のサービス提供企業のコンピュータで構築できるようになる。すなわち、既知組成式リスト 1 0 1 を保持する企業は、既知組成式リスト 1 0 1 を提供することなくサービス提供企業に対して材料情報出力装置 1 0 0 を用いた材料情報に関するサービスを実施させることができる。

【0112】

この態様を実現するには、図 1 において、材料情報出力装置 1 0 0 に含まれていない残りのブロックを材料情報出力装置 1 0 0 とは別の第 1 コンピュータで構成すればよい。そして、材料情報出力装置 1 0 0 とこの第 1 コンピュータとをインターネット等のネットワークを介して通信可能に接続すればよい。或いは、図 1 において入力部 1 0 3 及び表示部 1 1 0 を更に別の第 2 コンピュータで構成してもよい。この場合、第 2 コンピュータはサービス提供企業からサービスが提供されるユーザが保持するコンピュータである。これにより、サービス提供企業は、既知組成式リスト 1 0 1 を保持する企業以外のユーザに対してサービスを提供することができる。

10

20

30

40

50

【0113】

(実施の形態1の変形例1)

図13は、実施の形態1の変形例に係る材料情報出力システムの構成の一例を示すブロック図である。この変形例では、教師ラベル付与部105は、既知組成式リスト101に含まれる既知組成式を読み込む。教師ラベル付与部105は、既知組成式リスト101に含まれる既知組成式に対して合成可能性の最大値である「1」を教師ラベルLとして付与する。一方、教師ラベル付与部105は、疑似組成式に対しては、合成可能性の上限値を「 $1 - \varepsilon$ 」に設定して教師ラベルLを付与する。

【0114】

これにより、推定モデル学習部106は、教師ラベルLが付与された疑似組成式のみならず、教師ラベルLが付与された既知組成式も学習データとして利用して推定モデルの学習を行うことができる。なお、 ε の値としては、例えば $\varepsilon = 0.001$ などの値が設定されるが、これは一例であり、適宜変更可能である。具体的には、教師ラベル付与部105は、疑似組成式に対して算出した教師ラベルLに $1 - \varepsilon$ を乗じることで教師ラベルLを補正すればよい。

【0115】

これにより、確実に合成可能である既知組成式を考慮に入れて、推定モデルを学習させることができ、推定モデルの推定精度の向上を図ることができる。

【0116】

(実施の形態1の変形例2)

教師ラベル付与部105は、学習過程の初期に生成された疑似組成式、すなわち、学習段階を示す数値tが0の疑似組成式と、学習終了時に生成された疑似組成式、すなわち、学習段階を示す数値tがTの疑似組成式とにのみ教師ラベルLを付与し、これらの疑似組成式のみを疑似組成式保持部104に保持させてもよい。これにより、疑似組成式保持部104のメモリの空き容量を確保できる。

【0117】

(実施の形態2)

実施の形態2に係る材料情報出力システムは、合成可能性を推定したい材料の種類を考慮に入れて、推定モデルを生成するものである。図14は、本開示の実施の形態2に係る材料情報出力システムの構成の一例を示すブロック図である。なお、実施の形態2において、実施の形態1と同一の構成要素には同一の符号を付し、詳細な説明は省略する。

【0118】

図14において図1との相違点は、組み合わせ算出部1301と組み合わせ保持部1302とが新たに追加されている点にある。組み合わせ算出部1301は、入力部103で入力された評価対象となる材料の種類情報が示す種類と同じ種類の既知組成式を第1既知組成式として既知組成式リスト101から抽出する。そして、組み合わせ算出部1301は、後述する投票処理を実行し、得られた投票結果を組み合わせ保持部1302に記憶させる。

【0119】

組み合わせ保持部1302は、例えば、半導体メモリで構成され、組み合わせ算出部1301が生成した投票結果を保持する。

【0120】

教師ラベル付与部105は、疑似組成式保持部104が保持する疑似組成式と、組み合わせ保持部1302が保持する投票結果とを照らし合わせ、疑似組成式保持部104が保持する疑似組成式の中から第1疑似組成式を抽出する。そして、教師ラベル付与部105は、第1疑似組成式のそれぞれに対して実施の形態1と同一の手法を用いて教師ラベルを付与する。

【0121】

図15は、図14に示す材料情報出力システムの学習過程における処理の一例を示すフローチャートである。なお、図15のフローにおいて図10と同一の処理には同一のステ

10

20

30

40

50

ップ番号を付している。

【0122】

S301では、組み合わせ算出部1301は、既知組成式リスト101に含まれる既知組成式と、ユーザが入力した評価対象の材料の種類を示す種類情報とを取得する。ここで、種類情報とは、例えば、リチウムイオン電池及びマグネシウム二次電池等の材料を分類する情報である。なお、種類情報は、固体電池、太陽電池、及び熱電材料等の情報が採用されてもよい。

【0123】

以降、実施の形態1で説明したS102からS110までの処理とS302との処理とが並列に実行される。S302では、組み合わせ算出部1301は、投票処理を実行する。図16は、投票処理の詳細を示すフローチャートである。

10

【0124】

S1601では、組み合わせ算出部1301は、既知組成式リスト101からS301で取得された種類情報が示す種類と同一種類の既知組成式を第1既知組成式として抽出する。

【0125】

図18は、実施の形態2に係る既知組成式リスト101のデータ構成の一例を示す図である。図18においては、図2に対して更に「種類」が設けられている。「種類」は、「リチウムイオン電池、マグネシウム電池といった各組成式が示す材料の種類情報である。例えば、入力された種類情報がリチウムイオン電池であれば、「種類」がリチウムイオン電池である既知組成式が第1既知組成式として抽出される。

20

【0126】

図16に参照を戻す。S1602では、組み合わせ算出部1301は、第1既知組成式のそれぞれについて元素の族の組み合わせを特定する。

【0127】

元素の族とは、元素の周期表の縦1列に相当するクラスタのことを指すが、本実施の形態では、元素の族は、図17に示すように周期表を7つのクラスタに分類することで構成される。図17は、元素の族を説明する図である。図17に示すように、本実施の形態では、元素の族は、アルカリ金属1401と、アルカリ土類金属1402と、ハロゲン1403と、希ガス1404と、遷移元素1405と、その他の金属元素1406と、その他の非金属元素1407とで構成される。

30

【0128】

元素の族の組み合わせとは、二元系であれば、図16に示す7つの元素の族のうちの任意の2つの元素の族の組み合わせが該当し、三元系であれば、図16に示す7つの元素の族のうちの任意の3つの元素の族の組み合わせが該当する。したがって、二元系であれば、元素の族の組み合わせは $7C_2 = 21$ 通り存在し、三元系であれば、元素の族の組み合わせは $7C_3 = 35$ 通り存在する。

【0129】

例えば、 Al_2O_3 とMgOとは共に、元素の族の組み合わせが「その他の金属元素1406」と「その他の非金属元素1407」とであるため、元素の族の組み合わせは同一となる。一方、 $MgCl_2$ は、元素の族の組み合わせが「その他の金属元素1406」と「ハロゲン1403」とであるため、 Al_2O_3 とは元素の族の組み合わせが異なるものとなる。

40

【0130】

図16に参照を戻す。S1603では、組み合わせ算出部1301は、周期表から得られる予め定められた元素の族の組み合わせに対して、第1既知組成式の元素の族の組み合わせを投票する。

【0131】

例えば、第1既知組成式に Al_2O_3 が含まれているとすると、「その他の金属元素1406」と「その他の非金属元素1407」との元素の組み合わせに対して1票が投じら

50

れる。第1既知組成式に MgCl_2 が含まれているとすると、「その他の金属元素1406」と「ハロゲン1403」との元素の組み合わせに対して1票が投じられる。なお、周期表から得られる予め定められた元素の族の組み合わせとは、二元系であれば上述の21通りの元素の族の組み合わせが該当し、三元系であれば上述の70通りの元素の族の組み合わせが該当する。したがって、投票結果は、例えば、「その他の金属元素1406」と「ハロゲン1403」との元素の族の組み合わせについてa票、「その他の金属元素1406」と「その他の非金属元素1407」との元素の族の組み合わせについてb票というように、予め定められた元素の族の組み合わせと、それぞれの元素の族の組み合わせに対する投票数とが対応付けられたデータ構成を持つことになる。

【0132】

S1604では、組み合わせ算出部1301は、投票結果を組み合わせ保持部1302に保持させる。以上により投票処理が終了される。

【0133】

図15に参照を戻す。S303では、教師ラベル付与部105は、疑似組成式保持部104に保持されている疑似組成式のうち、元素の族の組み合わせが、組み合わせ保持部1302に保持されている投票結果が0でない元素の族の組み合わせを持つ疑似組成式を第1疑似組成式として抽出する。そして、教師ラベル付与部105は、第1疑似組成式に対して実施の形態1と同じ手法を用いて教師ラベルLを付与する。

【0134】

以後、S112、S113の処理が実施され、S303で教師ラベルLが付与された疑似組成式を用いて推定モデルが生成される。

【0135】

元素の族が異なると、元素の持つ特徴が大きく変化する。加えて、機械学習において、評価対象の材料に関連するデータのデータ数の比率が他の材料と比べて少ない場合、評価対象の材料の合成可能性の推定精度が低下する。そこで、本実施の形態では、評価対象の材料と種類が同じである第1既知組成式が抽出され、第1既知組成式と元素の族の組み合わせが同じである第1疑似組成式を用いて推定モデルが生成されている。そのため、本実施の形態では、評価対象の材料（組成式）の合成可能性の推定精度を向上させることができる。

【0136】

一方で、生成モデル学習部102における生成部802の学習においては、様々な種類のデータを利用した方が、これまでの経験では思いもつかない疑似組成式を生成する可能性が高まるため、データを限定すべきではない。そこで、本実施の形態では、実施の形態1と同様、生成モデル学習部102は既知組成式リスト101に含まれる全ての既知組成式を利用して生成モデルを生成する。

【0137】

以上説明したように、実施の形態2に係る材料情報出力システムによれば、評価対象の材料の種類をユーザに選択できるようにすることによって、評価対象の材料（組成式）の合成可能性の推定精度を向上させることができる。

【0138】

（実施の形態3）

実施の形態3に係る材料情報出力システムは、疑似組成式の形成エネルギーを推定し、推定結果に応じて疑似組成式に付与された教師ラベルLを補正するものである。

【0139】

形成エネルギーとは、物質が安定に存在しうるかに関わる特性値である。形成エネルギーが低すぎる場合は、物質の安定性が低下する。一方、形成エネルギーが高すぎる場合も物質の安定性は低下する。そのため、適度な形成エネルギーを持つことで、物質は安定に存在することができる。物質が安定に存在しうるということは、合成の容易さを意味する。そのため、本実施の形態においては、疑似組成式の形成エネルギーを推定し、推定結果を用いて教師ラベルLを補正する。

10

20

30

40

50

【0140】

図19は、本開示の実施の形態3に係る材料情報出力システムの構成の一例を示すブロック図である。なお、本実施の形態において、実施の形態1、2と同一の構成要素には同一の符号を付し、説明を省略する。

【0141】

図19において、図1との相違点は、エネルギー推定モデル学習部1601と、エネルギー推定モデル保持部1602とが新たに追加されている点にある。本実施の形態においては、既知組成式リスト101に含まれる既知組成式の特性値には、形成エネルギーが含まれているものとする。すなわち、本実施の形態では、既知組成式の形成エネルギーがメモリに事前に記憶されている。

【0142】

エネルギー推定モデル学習部1601は、既知組成式リスト101から各既知組成式について、特徴量及び記述子と形成エネルギーとのペアの情報を読み込み、形成エネルギーを推定するためのニューラルネットワークの荷重 W を調整することでエネルギー推定モデルを生成する。

【0143】

このとき、形成エネルギーは実数値となるため、このニューラルネットワークは回帰問題を解くこととなる。すなわち、エネルギー推定モデルは、特徴量及び記述子を入力データ X とし、形成エネルギーを出力データ Y とするニューラルネットワークである。

【0144】

そして、エネルギー推定モデル学習部1601は、学習の結果得られたエネルギー推定モデルに、既知組成式リスト101に含まれる各既知組成式の特徴量及び記述子を入力することで、各既知組成式に対する形成エネルギーを得る。そして、エネルギー推定モデル学習部1601は、得られた各既知組成式の形成エネルギーの平均と分散とを算出する。ここでは、既知エネルギーの平均と分散との算出にはエネルギー推定モデルが推定した形成エネルギーが用いられたが、既知組成式リスト101に記憶されている形成エネルギーが用いられてもよい。

【0145】

荷重 W の調整が完了したニューラルネットワークは、エネルギー推定モデルとしてエネルギー推定モデル保持部1602に保持される。エネルギー推定モデル保持部1602に保持される情報には、推定モデル保持部107と同様、エネルギー推定モデルを構成するニューラルネットワークが有する層の数及び各層に配置されるユニット数を表す構成情報と、各ユニットでの荷重和計算に用いられる荷重値を表す荷重 W とが含まれる。加えて、エネルギー推定モデル学習部1601によって算出された既知組成式の形成エネルギーの平均及び分散もエネルギー推定モデル保持部1602に保持される。

【0146】

教師ラベル付与部105は、実施の形態1の手法を用いて疑似組成式に教師ラベル L を付与すると共に、エネルギー推定モデルにより推定された疑似組成式の形成エネルギーを用いて教師ラベル L を補正する。

【0147】

具体的には、まず、教師ラベル付与部105は、実施の形態1の方法で疑似組成式保持部104に保持された各疑似組成式に教師ラベル L を付与する。次に、教師ラベル付与部105は、エネルギー推定モデル保持部1602からエネルギー形成モデルを取得する。次に、教師ラベル付与部105は、疑似組成式保持部104に保持された各疑似組成式の特徴量及び記述子をエネルギー形成モデルに入力することで、各疑似組成式の形成エネルギーを得る。次に、教師ラベル付与部105は、エネルギー推定モデル保持部1602に保持されている既知組成式の形成エネルギーの平均及び分散を読み込む。次に、教師ラベル付与部105は、読み込んだ平均及び分散と、各疑似組成式の形成エネルギーとを用いて、各疑似組成式のマハラノビス距離 M を求める。

【0148】

次に、教師ラベル付与部 105 は、疑似組成式に付与されている教師ラベル L に対して $L' = \exp(-M) \times L$ の演算を行い、教師ラベル L を補正する。 L' は補正後の教師ラベルを表す。これは、学習回数に基づいて得られた教師ラベル L に形成エネルギーに応じた重み付けを行ったことを意味する。これにより、教師ラベル L は、マハラノビス距離 M が増大するにつれて値が小さくなるように補正されるのである。

【0149】

上述したように、形成エネルギーが値が高すぎる又は低すぎると、物質は安定しない。そのため、既知組成式の形成エネルギーに対してかけ離れた形成エネルギーを持つ疑似組成式は合成可能性が低いと言える。そこで、本実施の形態では、既知組成式の形成エネルギーの集団に対して疑似組成式の形成エネルギーがどの程度離れているかを示すマハラノビス距離 M を用いて教師ラベル L を補正する。これにより、本実施の形態では、既知組成式に対して物理学の知見を踏まえた教師ラベル L' を付与することができる。その結果、本実施の形態は、合成可能性が高精度に推定できる推定モデルを生成することができる。

10

【0150】

なお、本実施の形態では距離としてマハラノビス距離 M が採用されたがこれは一例であり、既知組成式の形成エネルギーの平均に対するユークリッド距離等、既知組成式の形成エネルギーの集団に対する距離が測定可能な距離であればどのような距離が採用されてもよい。

【0151】

(実施の形態 4)

20

実施の形態 4 に係る材料情報出力システムは、特性値の範囲をユーザが入力すると、その範囲内の特性値を持つ疑似組成式を合成可能性と合わせてユーザに提示するものである。

【0152】

図 20 は、本開示の実施の形態 4 に係る材料情報出力システムの構成の一例を示すブロック図である。なお、本実施の形態において実施の形態 1～3 と同一の構成要素には同一の符号を付し説明を省略する。

【0153】

図 20 において、図 1 との相違点は、更に、特性値推定モデル保持部 1701 と、生成モデル保持部 1702 と、特性値算出部 1703 とが材料情報出力装置 100 に設けられている点にある。

30

【0154】

実施の形態 1 で示したように、生成モデル学習部 102 の学習では、生成部 802 と識別部 803 とによる敵対的学習アルゴリズムが用いられている。学習の結果得られた生成部 802 のニューラルネットワーク NN1 は、特性値と乱数の発生数とが設定されると、設定された特性値を持つことが推定される疑似組成式を、設定した乱数の発生数だけ生成できる。また、識別部 803 のニューラルネットワーク NN2 は、疑似組成式の特徴量及び記述子が入力されると、その疑似組成式の特性値の推定結果を出力できる。

【0155】

そこで、本実施の形態では、学習済みのニューラルネットワーク NN1 と学習済みのニューラルネットワーク NN2 とを活用過程で利用する。

40

【0156】

すなわち、特性値推定モデル保持部 1701 は、学習済みのニューラルネットワーク NN2 を特性値推定モデルとして保持する。生成モデル保持部 1702 は、学習済みのニューラルネットワーク NN1 を生成モデルとして保持する。

【0157】

活用過程において、入力部 103 は、組成式ではなく、特性値の範囲がユーザにより入力される。特性値は、例えば、電池の充放電率又は容量等である。或いは、特性値は、ヤング率又はイオン化ポテンシャル等であってもよい。ユーザである材料開発者は、所望の特性値の範囲を満足する組成式を探索したい場合がある。そこで、本実施の形態では、ユーザに入力部 103 を用いて特性値の範囲を入力させる。加えて、ユーザに入力部 103 を

50

用いて生成したい疑似組成式の生成数Nを入力させる。ここで生成数Nは、表示部110での視認が容易な数であり、例えば10や100である。

【0158】

記述子算出部108は、生成モデル保持部1702から生成モデルを取得し、その生成モデルに対して特性値を入力し、特徴量及び記述子を算出する。

【0159】

具体的には、記述子算出部108は、ユーザが入力した特性値の範囲内の特性値と乱数とのペアからなる入力データ $X = [\text{乱数}, \text{特性値}]$ をN個生成する。ここで、N個の入力データXは、 $X1 = [\text{乱数}A1, \text{特性値}B1]$ 、 $X2 = [\text{乱数}A2, \text{特性値}B2]$ 、 $\dots$ 、 $XN = [\text{乱数}AN, \text{特性値}BN]$ と表される。

【0160】

そして、記述子算出部108は、生成モデルにN個の入力データXを順次入力し、特徴量及び記述子を表すN個の出力データY、すなわち、N個の疑似組成式を得る。ここで、記述子算出部108は、ユーザが入力部103に入力した特性値の範囲内からランダムにN個の値をサンプリングすることでN個の特性値を生成すればよい。

【0161】

特性値算出部1703は、特性値推定モデル保持部1701に保持されている特性値推定モデルを取得する。そして、特性値算出部1703は、記述子算出部108で得られたN個の特徴量及び記述子を入力データXとして特性値推定モデルに入力し、N個の疑似組成式に対する特性値の推定結果を表すN個の出力データYを算出する。

【0162】

合成可能性算出部109は、推定モデル保持部107から推定モデルを取得し、記述子算出部108が算出したN個の特徴量及び記述子を推定モデルに入力し、N個の疑似組成式に対する合成可能性を算出する。

【0163】

表示部110は、合成可能性を示す第1軸と、特性値を示す第2軸とを備えるグラフ上に、N個の疑似組成式についての合成可能性と特性値とがプロットされた散布図を生成して表示する。

【0164】

図21は、表示部110が表示する散布図G2の一例を示す図である。散布図G2は、縦軸を特性値、横軸を合成可能性とする2軸のグラフ1801を備える。グラフ1801の下には、ユーザが入力部103を用いて入力した特性値の範囲と、疑似組成式の生成数とが表示されている。ここでは、特性値として50～500、生成数として6が入力されている。

【0165】

グラフ1801には、生成された6個の疑似組成式を示す6個の点PP1がプロットされている。また、各点PP1の近傍には、特徴量及び記述子から算出される組成式が表示されている。実施の形態1でも示したように、特徴量及び記述子から組成式への変換は容易である。この表示によって、材料開発者は、狙いとする特性値の範囲内でどのような組成式が生成可能であるかを容易に認識できる。

【0166】

図22は、本実施の形態4に係る材料情報出力システムの活用過程における処理の一例を示すフローチャートである。

【0167】

S401では、入力部103は、ユーザにより、特性値の範囲と生成したい疑似組成式の生成数Nとが入力される。S402では、記述子算出部108は、入力された特性値の範囲から特性値と乱数とのペアからなるN個の入力データXを生成する。S403では、記述子算出部108は、生成モデル保持部1702から生成モデルを取得し、生成モデルにN個の入力データ $X = [\text{乱数}, \text{特性値}]$ を入力し、N個の疑似組成式を生成する。

【0168】

10

20

30

40

50

S 4 0 4では、特性値算出部 1 7 0 3は、特性値推定モデル保持部 1 7 0 1から特性値推定モデルを取得し、N個の疑似組成式を特性値推定モデルに入力し、N個の特性値を算出する。

【0169】

S 4 0 4と同時に、S 4 0 5では、合成可能性算出部 1 0 9は、推定モデル保持部 1 0 7から推定モデルを取得し、推定モデルにN個の疑似組成式を入力し、N個の疑似組成式のそれぞれに対する合成可能性を算出する。

【0170】

S 4 0 6では、表示部 1 1 0はS 4 0 3で算出された疑似組成式の特徴量及び記述子から対応する組成式を算出し、N個の疑似組成式をグラフ 1 8 0 1にプロットすると共に、プロットした各点 P P 1の近傍に対応する組成式がグラフ 1 8 0 1上に表示された散布図 G 2を生成して表示する。

【0171】

以上説明したように、実施の形態 4に係る材料情報出力システムによれば、ユーザが狙いとする特性値の範囲を入力すると、その範囲内の特性値を持つ疑似組成式が生成され、その疑似組成式がグラフ 1 8 0 1上にプロットされた散布図 G 2が生成される。そのため、本実施の形態は、材料開発者の解析を補助するための有用な判断材料を提示できる。

【0172】

なお、実施の形態 4では、グラフ 1 8 0 1は2軸で構成されているがこれは一例であり、特性値及び合成可能性以外に表示したいパラメータがあれば、3軸以上のグラフで構成されてもよい。

【0173】

(ハードウェア構成)

材料情報出力システムは、コンピュータを利用して実現される。図 2 3は、材料情報出力システムを実現するためのハードウェア構成の一例を示すブロック図である。材料情報出力システムは、コンピュータ 2 0 0 0と、コンピュータ 2 0 0 0に指示を与えるためのキーボード 2 0 1 1及びマウス 2 0 1 2と、コンピュータ 2 0 0 0の演算結果等の情報を提示するためのディスプレイ 2 0 1 0と、コンピュータ 2 0 0 0で実行されるプログラムを読み取るためのODD (Optical Disk Drive) 2 0 0 8とを含む。

【0174】

材料情報出力システムが実行するプログラムは、コンピュータで読み取り可能な光記憶媒体 2 0 0 9に記憶され、ODD 2 0 0 8で読み取られる。または、コンピュータネットワークを通じてNIC 2 0 0 6で読み取られる。

【0175】

コンピュータ 2 0 0 0は、CPU (Central Processing Unit) 2 0 0 1と、ROM (Read Only Memory) 2 0 0 4と、RAM (Random Access Memory) 2 0 0 3と、HDD (Hard Disk Drive) 2 0 0 5と、NIC (Network Interface Controller) 2 0 0 6と、バス 2 0 0 7とを含む。

【0176】

更に、コンピュータ 2 0 0 0は、高速演算を行うためのGPU (Graphical Processing Unit) 2 0 0 2を含んでも良い。

【0177】

CPU 2 0 0 1とGPU 2 0 0 2とは、ODD 2 0 0 8またはNIC 2 0 0 6を介して読み取られたプログラムを実行する。ROM 2 0 0 4は、コンピュータ 2 0 0 0の動作に必要なプログラム及びデータを記憶する。RAM 2 0 0 3は、プログラム実行時のパラメータなどのデータを記憶する。HDD 2 0 0 5は、プログラム及びデータなどを記憶する。NIC 2 0 0 6は、コンピュータネットワークを介して他のコンピュータとの通信を行う。バス 2 0 0 7は、CPU 2 0 0 1、ROM 2 0 0 4、RAM 2 0 0 3、HDD 2 0 0 5、NIC 2 0 0 6、ディスプレイ 2 0 1 0、キーボード 2 0 1 1、マウス 2 0 1 2及び

ODD2008を相互に接続する。

【0178】

尚、コンピュータ2000に接続されているキーボード2111、マウス2012、及びODD2008は、例えばディスプレイ2010がタッチパネルになっている場合又はNIC2006を利用する場合には、取り外しても良い。

【0179】

図23に示すコンピュータ2000は、図1に示す材料情報出力装置100と上述した第1コンピュータと、第2コンピュータとのそれぞれを構成してもよい。

【0180】

更に、上記の各装置を構成する材料情報出力システムの構成要素の一部または全ては、1個のシステムLSI (Large Scale Integration: 大規模集積回路) から構成されてもよい。システムLSIは、複数の構成部を1個のチップ上に蓄積して製造された超多機能LSIであり、具体的には、マイクロプロセッサ、ROM、RAMなどを含んで構成されるコンピュータシステムである。RAMには、コンピュータプログラムが記憶されている。マイクロプロセッサが、コンピュータプログラムに従って動作することにより、システムLSIは、その機能を達成する。

【0181】

また、上記の各装置を構成する構成要素の一部または全ては、各装置に着脱可能なICカードまたは単体モジュールから構成されてもよい。ICカードまたはモジュールは、マイクロプロセッサ、ROM、RAMなどから構成されるコンピュータシステムである。ICカードまたはモジュールは、上記の超多機能LSIを含むとしても良い。マイクロプロセッサが、コンピュータプログラムに従って動作することにより、ICカードまたはモジュールは、その機能を達成する。このICカードまたはこのモジュールは、耐タンパ性を有するとしても良い。

【0182】

また、本開示は、上記に示す材料情報出力システムの処理を実行する方法であるとしても良い。また、これらの方法をコンピュータにより実現するコンピュータプログラムを含んでも良いし、前記コンピュータプログラムからなるデジタル信号を含んでも良い。

【0183】

更に、本開示は、上記コンピュータプログラム又は上記デジタル信号をコンピュータ読み取り可能な非一時的な記憶媒体、例えば、フレキシブルディスク、ハードディスク、CD-ROM、MO、DVD、DVD-ROM、DVD-RAM、BD (Blue-ray (登録商標) Disc)、半導体メモリなどに記憶したものを含んでも良い。また、これら非一時的な記憶媒体に記録されている上記デジタル信号を含んでも良い。

【0184】

また、本発明は、上記コンピュータプログラムまたは上記デジタル信号を、電気通信回線、無線または有線通信回路、インターネットを代表とするネットワーク、データ放送等を経由して伝送するものとしても良い。

【0185】

また、上記プログラムまたは上記デジタル信号を上記非一時的な記憶媒体に記録して移送することにより、または上記プログラムまたは上記デジタル信号は上記ネットワーク等を経由して移送することにより、独立した他のコンピュータシステムにより実施するとしても良い。

【0186】

図24は、本開示の材料情報出力システムのハードウェア構成の他の一例を示す図である。図24の例では、コンピュータ2000と、コンピュータ2000とネットワークを介して接続されたデータサーバ2101とを備えている。データサーバ2101は、データを記憶するメモリを備える。このデータは、上記ネットワーク等を経由してコンピュータ2000に読み出される。また、データサーバ2101から情報を読み出す、コンピュータ2000は1台である必要はなく、複数であっても良い。その際、各コンピュータ2

10

20

30

40

50

00が材料情報出力システムの構成要素の一部をそれぞれ実施しても良い。

【0187】

図24のハードウェア構成を採用する場合、推定モデル保持部107、疑似組成式保持部104、既知組成式リスト101をデータサーバ2101に設ければよい。また、図24のハードウェア構成を採用する場合、図1等を示す各ブロックは、データサーバ2101及びコンピュータ2000に分散して実装されてもよい。

【0188】

更に、上記実施の形態および上記変形例をそれぞれ組み合わせるとしても良い。

【0189】

今回開示された実施の形態は全ての点で例示であって制限的なものではないと考えられるべきである。本開示の範囲は上記した説明ではなく、請求の範囲によって示され、請求の範囲と均等の意味および範囲内での全ての変更が含まれることが意図される。

【産業上の利用可能性】

【0190】

本発明によると、未知の組成式に対する適切な合成可能性を提示できるので、新規材料開発の分野にとって有用である。

【符号の説明】

【0191】

100 : 材料情報出力装置

101 : 既知組成式リスト

102 : 生成モデル学習部

103 : 入力部

104 : 疑似組成式保持部

105 : 教師ラベル付与部

106 : 推定モデル学習部

107 : 推定モデル保持部

108 : 記述子算出部

109 : 合成可能性算出部

110 : 表示部

801 : 乱数発生部

802 : 生成部

803 : 識別部

1301 : 組み合わせ算出部

1302 : 組み合わせ保持部

1601 : エネルギー推定モデル学習部

1602 : エネルギー推定モデル保持部

1701 : 特性値推定モデル保持部

1702 : 生成モデル保持部

1703 : 特性値算出部

10

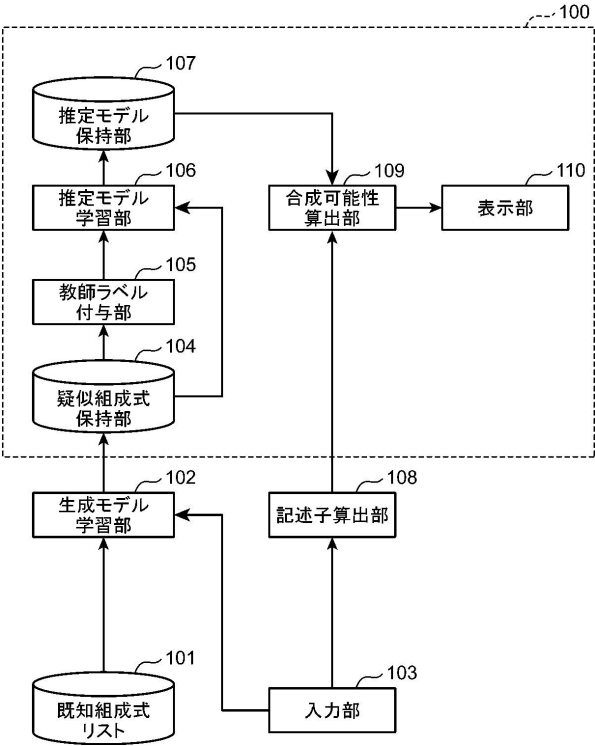
20

30

40

50

【図面】
【図 1】



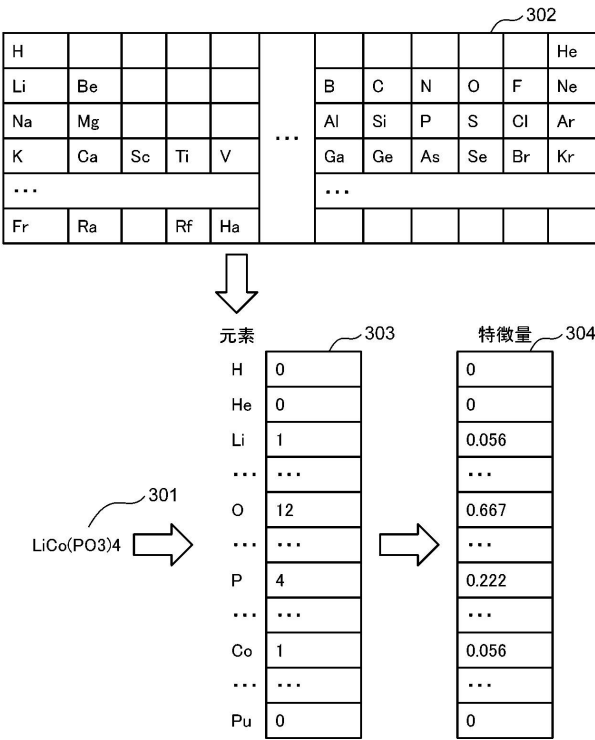
【図 2】

組成式	特性値		特徴量				記述子		
	電圧	容量					a	γ	
Li _{0.1} TiO ₂	9.43	1214	...	0	0.2	...	2.877	...	120
Li _{0.1} Fe ₅ O ₈	8.76	291		0	0.2		8.488		90
Li _{0.67-1} FeBO ₃	7.37	242		0	0.2		8.238		120
...									
Li _{0.1} MnF ₄	6.93	641		0	0.2		11.279		90

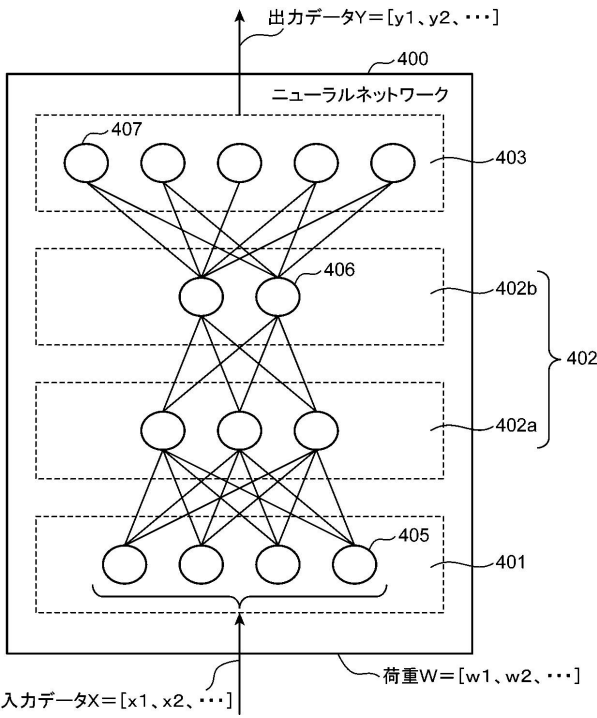
10

20

【図 3】



【図 4】

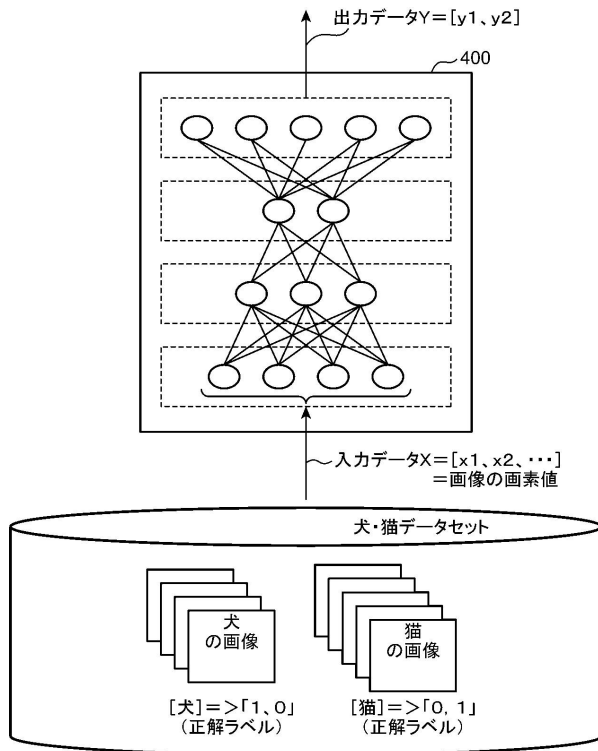


30

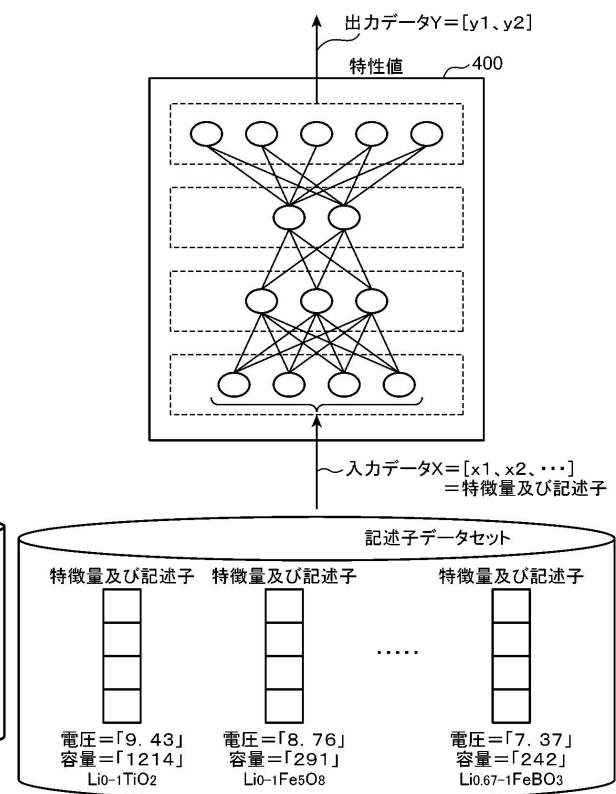
40

50

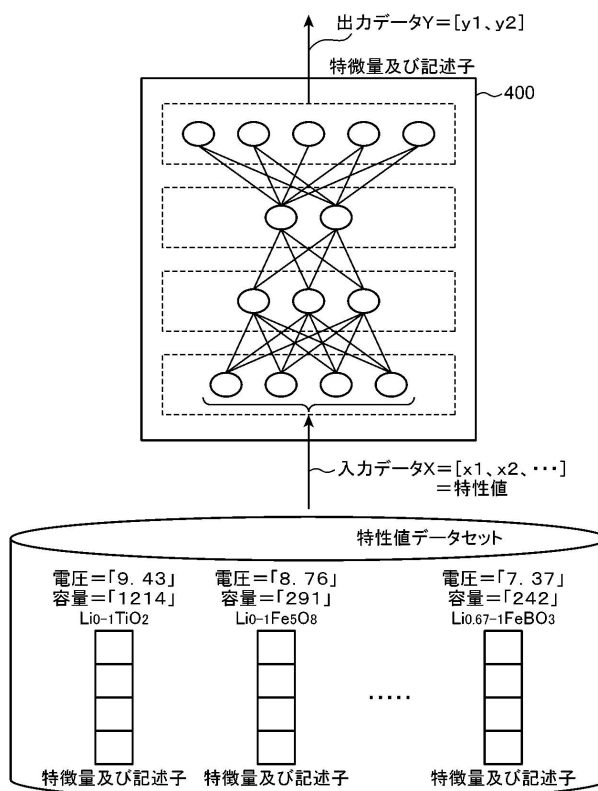
【図 5】



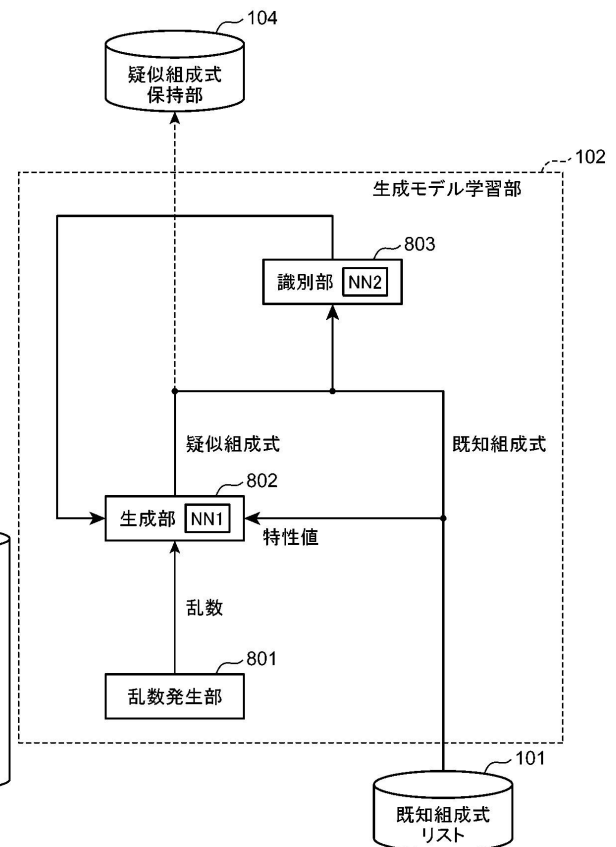
【図 6】



【図 7】



【図 8】



10

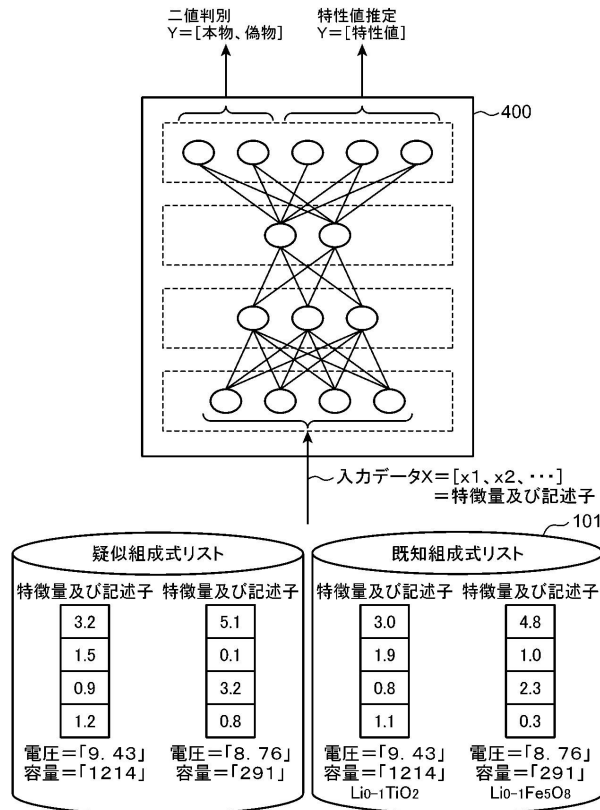
20

30

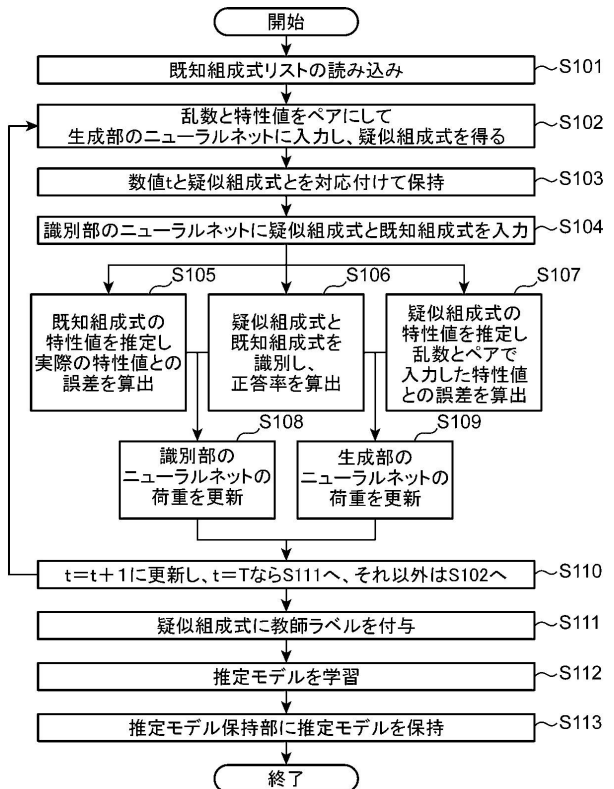
40

50

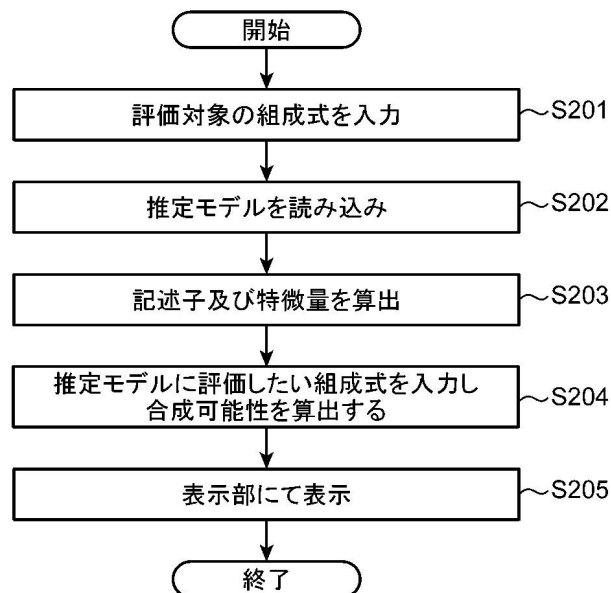
【図 9】



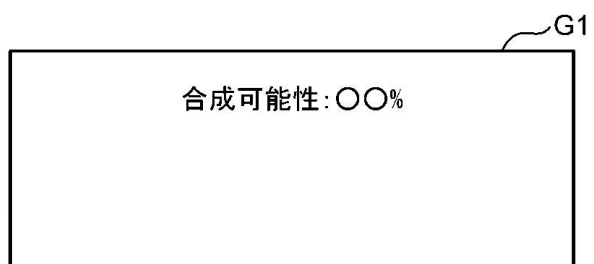
【図 10】



【図 11】



【図 12】



10

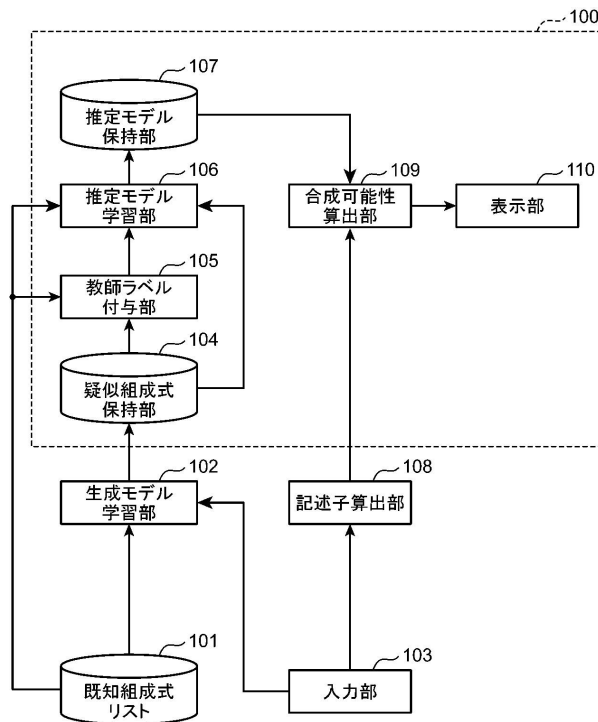
20

30

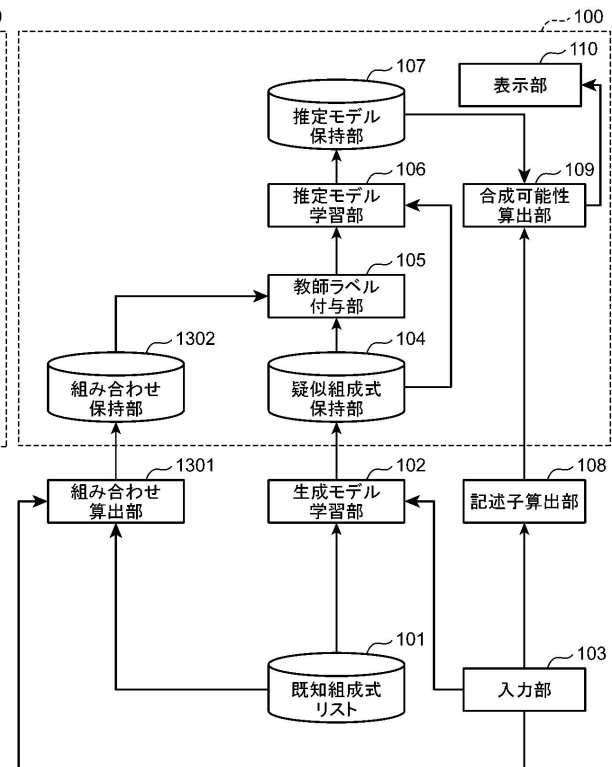
40

50

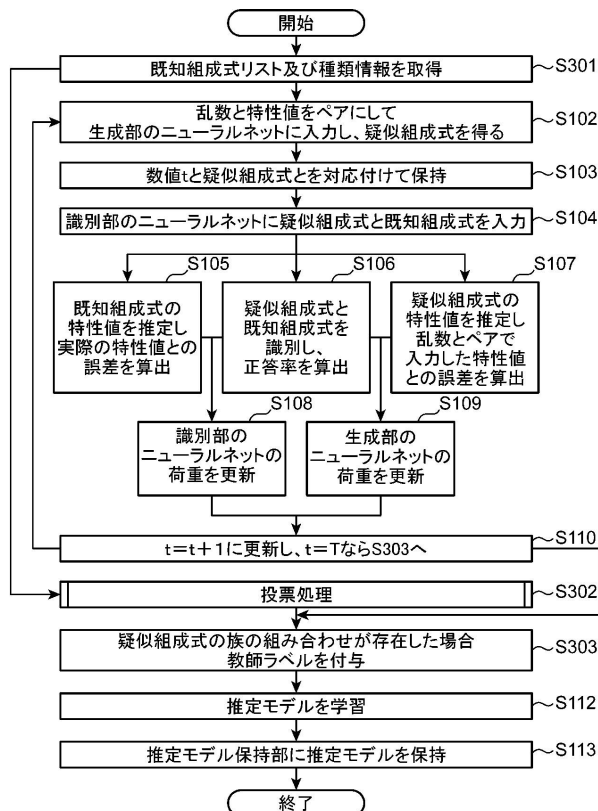
【図 1 3】



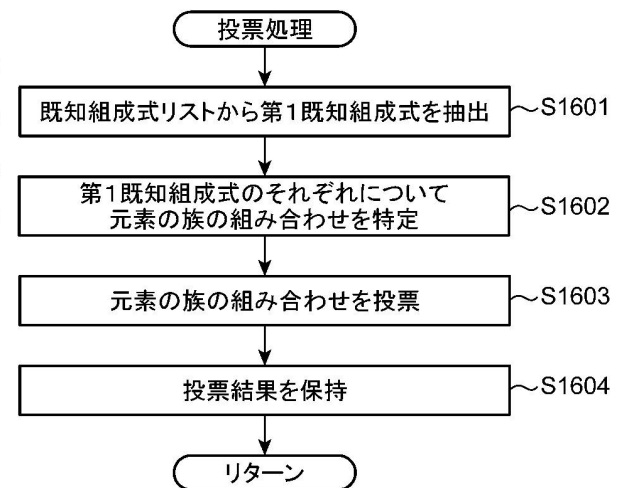
【図 1 4】



【図 1 5】



【図 1 6】



10

20

30

40

50

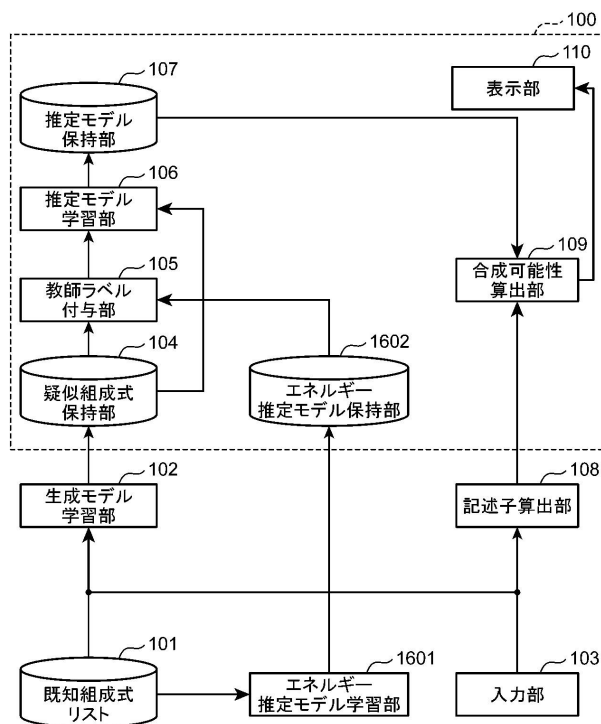
【図 17】

1406	H	1407																He		1404
																		Ne	Ar	
1402	Li	Be	Mg	B	C	N	O	F	1403											
	Na	Al		Si	P	S	Cl	Br		At										
1401	K	Ca	Sc	Ti	V	Cr	Mn	Fe	Co	Ni	Cu	Zn	Ga	Ge	As	Se	Br	1405		
	Rb	Sr	Y	Zr	Nb	Mo	Tc	Ru	Rh	Pd	Ag	Cd	In	Sn	Sb	Te	I			
1400	Cs	Ba	*	Hf	Ta	W	Re	Os	Ir	Pt	Au	Hg	Tl	Pb	Bi	Po	At	1403		
	Fr	Ra	**	Rf	Db	Sg	Bh	Hs	Mt	Ds	Rg	1405								

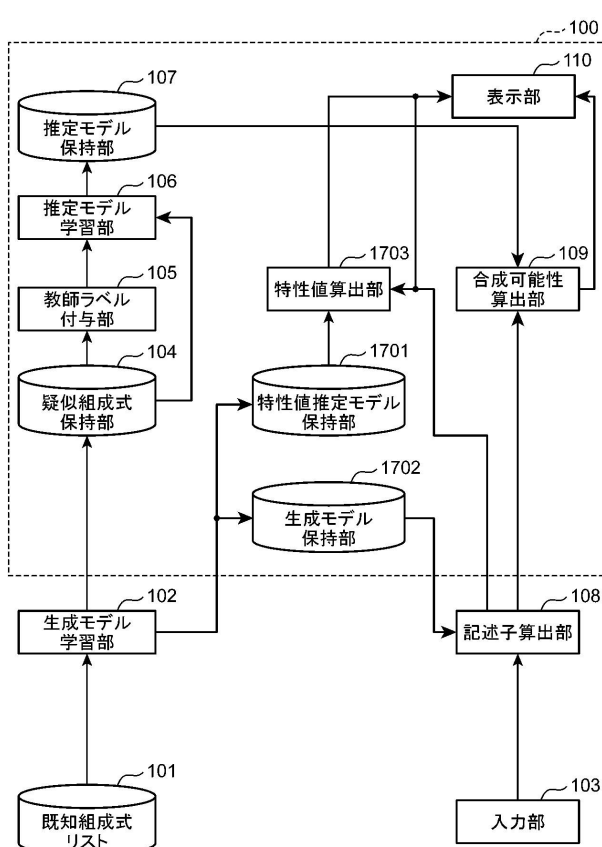
【図 18】

種類	組成式	特性値		特徴量			記述子	
		電圧	容量	...	0	0.2	a	γ
リチウムイオン電池	LiO-TiO_2	9.43	1214	...	0	0.2	2.877	120
マグネシウム電池	MgO-NiF_4	5.48	1384		0	0	5.055	90
アルミニウム電池	$\text{AlO}_{33-1}\text{FeF}_5$	4.28	1031		0	0	3.608	90
...	...							
リチウムイオン電池	LiO-MnF_4	6.93	641		0	0.2	11.279	90

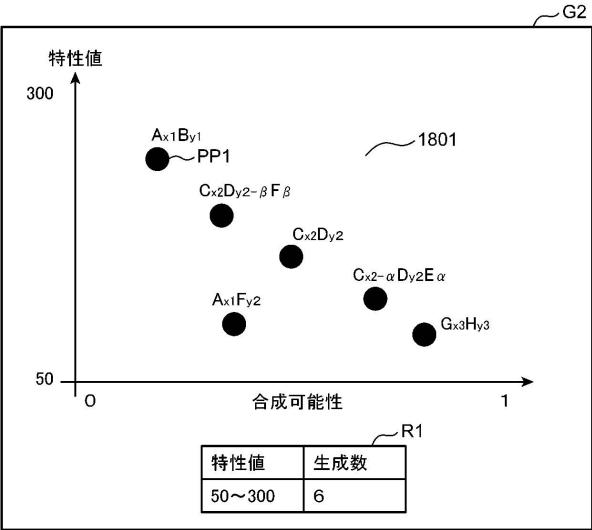
【図 19】



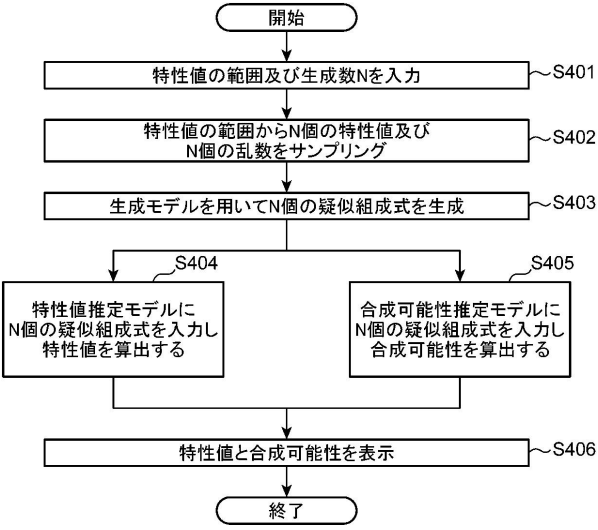
【図 20】



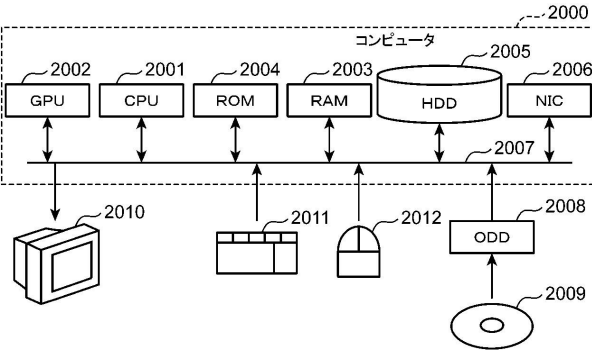
【図 2 1】



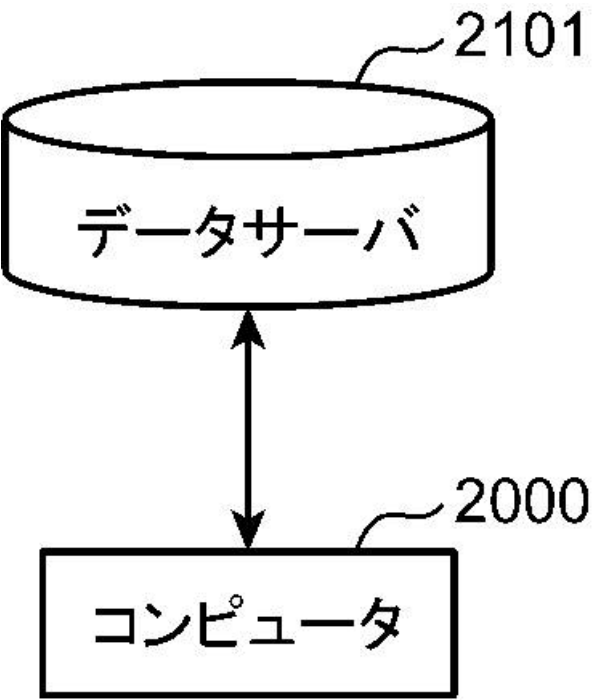
【図 2 2】



【図 2 3】



【図 2 4】



10

20

30

40

50

フロントページの続き

ナソニック株式会社内

(72)発明者 中田 透

大阪府門真市大字門真１００６番地 パナソニック株式会社内

審査官 池田 聡史

(56)参考文献 国際公開第２００８／０３５４１９（WO，A１）

国際公開第２０１０／０１６１０９（WO，A１）

特開２０１３－１９０２１６（JP，A）

特開２０１１－０７６９７５（JP，A）

特開平０１－１６１５７８（JP，A）

特開２０１０－２６２９３０（JP，A）

特開２０１８－１２３０１７（JP，A）

特開平０７－０３６８４６（JP，A）

特開２００４－２７２２８０（JP，A）

(58)調査した分野 (Int.Cl.，DB名)

G 0 6 Q 1 0 / 0 0 - 9 9 / 0 0

G 1 6 C 6 0 / 0 0