

(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号
特許第7290133号
(P7290133)

(45)発行日 令和5年6月13日(2023.6.13)

(24)登録日 令和5年6月5日(2023.6.5)

(51)国際特許分類 F I
G 0 6 N 3/08 (2023.01) G 0 6 N 3/08

請求項の数 9 (全14頁)

(21)出願番号	特願2020-55576(P2020-55576)	(73)特許権者	000004260 株式会社デンソー 愛知県刈谷市昭和町1丁目1番地
(22)出願日	令和2年3月26日(2020.3.26)	(74)代理人	100113549 弁理士 鈴木 守
(65)公開番号	特開2021-157379(P2021-157379 A)	(74)代理人	100115808 弁理士 加藤 真司
(43)公開日	令和3年10月7日(2021.10.7)	(74)代理人	100169199 弁理士 石本 貴幸
審査請求日	令和4年5月19日(2022.5.19)	(72)発明者	三木 彰 愛知県刈谷市昭和町1丁目1番地 株式 会社デンソー内
		(72)発明者	門脇 正史 愛知県刈谷市昭和町1丁目1番地 株式 会社デンソー内

最終頁に続く

(54)【発明の名称】 情報処理装置、重要ノード特定方法、及び重要ノード特定プログラム

(57)【特許請求の範囲】

【請求項1】

入力情報に対する出力情報を出力する学習モデルを構成する複数のノードのうち、前記学習モデルから除外する前記ノードを選択する選択手段(20)と、

前記選択手段によって選択された前記ノードが除外された前記学習モデルの適切度の低下に基づいて、前記学習モデルを構成する前記複数のノードの中から、相対的に重要とされる前記ノードを特定する特定手段(30)と、

前記選択手段によって選択された前記ノードが除外された前記学習モデルに評価データを適用し、前記学習モデルの前記適切度を算出する適切度算出手段(24)と、

前記ノードが除外された前記学習モデルを構成する前記ノードの組み合わせと前記適切度との関係を表す回帰式を生成する回帰式生成手段(26)と、

前記ノードが除外された前記学習モデルを用いて、前記適切度算出手段及び前記回帰式生成手段での処理を少なくとも1回以上行う繰返制御手段(28)と、

を備え、

前記選択手段は、前記回帰式が前記適切度をより正確に予測するように前記ノードを選択し、

前記特定手段は、所定の終了条件を満たした時点で、前記適切度がより低下した前記学習モデルに含まれない前記ノードを、重要とされる前記ノードとして特定する、

情報処理装置(10)。

【請求項2】

10

20

前記選択手段は、前記学習モデルから除外する前記ノードを複数選択する、請求項 1 記載の情報処理装置。

【請求項 3】

前記特定手段は、前記学習モデルを構成する複数の前記ノードの探索結果に基づいて、相対的に重要とされる前記ノードを特定する、請求項 1 又は請求項 2 記載の情報処理装置。

【請求項 4】

前記回帰式生成手段は、前記適切度を z とし、前記学習モデルに含まれる前記ノードの選択識別子を s_i , s_j とした下記 (1) 式で示される回帰式を生成する、請求項 1 から請求項 3 の何れか 1 項記載の情報処理装置。

【数 1】

$$z = \sum_{i,j} a_{ij}s_i s_j + \sum_i b_i s_i \quad \dots (1)$$

a_{ij} , b_i : 重みパラメータ

【請求項 5】

前記回帰式生成手段は、前記回帰式の重みパラメータを確率分布として求める、請求項 4 記載の情報処理装置。

【請求項 6】

前記回帰式は、3 次以上の項を含む、請求項 1 から請求項 5 の何れか 1 項記載の情報処理装置。

【請求項 7】

前記情報処理装置の全てまたは一部分は、量子コンピュータである、請求項 1 から請求項 6 の何れか 1 項記載の情報処理装置。

【請求項 8】

入力情報に対する出力情報を出力する学習モデルを構成する複数のノードのうち、前記学習モデルから除外する前記ノードを選択する選択工程と、

前記選択工程によって選択された前記ノードが除外された前記学習モデルの適切度の低下に基づいて、前記学習モデルを構成する前記複数のノードの中から、相対的に重要とされる前記ノードを特定する特定工程と、

前記選択工程によって選択された前記ノードが除外された前記学習モデルに評価データを適用し、前記学習モデルの前記適切度を算出する適切度算出工程と、

前記ノードが除外された前記学習モデルを構成する前記ノードの組み合わせと前記適切度との関係を表す回帰式を生成する回帰式生成工程と、

前記ノードが除外された前記学習モデルを用いて、前記適切度算出工程及び前記回帰式生成工程での処理を少なくとも 1 回以上行う繰返制御工程と、

を有し、

前記選択工程は、前記回帰式が前記適切度をより正確に予測するように前記ノードを選択し、

前記特定工程は、所定の終了条件を満たした時点で、前記適切度がより低下した前記学習モデルに含まれない前記ノードを、重要とされる前記ノードとして特定する、重要ノード特定方法。

【請求項 9】

情報処理装置が備えるコンピュータを、

入力情報に対する出力情報を出力する学習モデルを構成する複数のノードのうち、前記学習モデルから除外する前記ノードを選択する選択手段と、

前記選択手段によって選択された前記ノードが除外された前記学習モデルの適切度の低下に基づいて、前記学習モデルを構成する前記複数のノードの中から、相対的に重要とされる前記ノードを特定する特定手段と、

前記選択手段によって選択された前記ノードが除外された前記学習モデルに評価データを

10

20

30

40

50

適用し、前記学習モデルの前記適切度を算出する適切度算出手段(24)と、
 前記ノードが除外された前記学習モデルを構成する前記ノードの組み合わせと前記適切度
 との関係を表す回帰式を生成する回帰式生成手段(26)と、
 前記ノードが除外された前記学習モデルを用いて、前記適切度算出手段及び前記回帰式生
 成手段での処理を少なくとも1回以上行う繰返制御手段(28)と、
 して機能させ、
 前記選択手段は、前記回帰式が前記適切度をより正確に予測するように前記ノードを選択
 し、
 前記特定手段は、所定の終了条件を満たした時点で、前記適切度がより低下した前記学習
 モデルに含まれない前記ノードを、重要とされる前記ノードとして特定する、
 重要ノード特定プログラム。

10

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、情報処理装置、重要ノード特定方法、及び重要ノード特定プログラムに關する。

【背景技術】

【0002】

従来から、学習モデルとしてニューラルネットワークを用いた機械学習が広く行われて
 いる。ここで、ニューラルネットワークは、複数のノード(ニューロンやフィルタともい
 う。)で構成される1又は複数の中間層が入力層と出力層との間に設けられることで構成
 される。

20

【0003】

特許文献1に記載のように、ニューラルネットワークのような数理的モデルはブラック
 ボックスモデルとも呼ばれ、ブラックボックスモデルが持つパラメータは通常物理的の内見
 や化学的な意味を持たないため、物理量や化学量との対応付けが難しい。

【先行技術文献】

【特許文献】

【0004】

【文献】特開2005-346714号公報

30

【発明の概要】

【発明が解決しようとする課題】

【0005】

しかしながら、使用される学習モデルがブラックボックスであることが好ましくなく、
 当該学習モデルを理解できるようにホワイトボックス化することが望まれる場合もある。

【0006】

本発明は上記背景に鑑み、複数のノードにより構成される学習モデルをホワイトボッ
 クス化することができる、情報処理装置、重要ノード特定方法、及び重要ノード特定プロ
 グラム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】

40

【0007】

本発明は上記課題を解決するために以下の技術的手段を採用する。特許請求の範囲及び
 この項に記載した括弧内の符号は、ひとつの態様として後述する実施形態に記載の具体的
 手段との対応関係を示す一例であって、本発明の技術的範囲を限定するものではない。

【0008】

本発明の一態様の情報処理装置(10)は、入力情報に対する出力情報を出力する学習
 モデルを構成する複数のノードのうち、前記学習モデルから除外する前記ノードを選択す
 る選択手段(20)と、前記選択手段によって選択された前記ノードが除外された前記学
 習モデルの適切度に基づいて、前記学習モデルを構成する前記複数のノードの中から、相
 対的に重要とされる前記ノードを特定する特定手段(30)と、を備える。

50

【発明の効果】

【0009】

本発明によれば、複数のノードにより構成される学習モデルをホワイトボックス化することができる。

【図面の簡単な説明】

【0010】

【図1】本実施形態の情報処理装置の構成を示す概略構成図である。

【図2】本実施形態のニューラルネットワークの構成を示す模式図である。

【図3】本実施形態の一つのノードを除外した場合における学習モデルのモデル精度を示す試験結果である。

【図4】本実施形態の二つのノードを除外した場合における学習モデルのモデル精度を示す試験結果である。

【図5】本実施形態のノード除外学習モデルにおけるノードの組み合わせとノード除外学習モデルのモデル精度とを示す模式図である。

【図6】本実施形態の重みパラメータの確率分布を示す模式図である。

【図7】本実施形態の重要ノード特定処理の流れを示すフローチャートである。

【発明を実施するための形態】

【0011】

以下、図面を参照して本発明の実施形態を説明する。なお、以下に説明する実施形態は、本発明を実施する場合の一例を示すものであって、本発明を以下に説明する具体的構成に限定するものではない。本発明の実施にあたっては、実施形態に応じた具体的構成が適宜採用されてよい。

【0012】

図1は、本実施形態の情報処理装置10の電気的構成を示すブロック図である。本実施形態に係る情報処理装置10は、量子コンピュータ等で構成される演算部12、各種プログラム及び各種データ等が予め記憶されたROM(Read Only Memory)14、演算部12による各種プログラムの実行時のワークエリア等として用いられるRAM(Random Access Memory)16、各種プログラムや各種データ等を記憶する大容量記憶装置18を備えている。

【0013】

なお、量子コンピュータとは、電子や光子などの量子力学の原理が妥当する物質を利用して、情報の基本単位(これを量子ビットという)を構成し、この量子ビットを用いて計算を行うコンピュータをいう。また、演算部12は、量子コンピュータに限らず、古典ビット(例えばCMOSトランジスタ)により演算を行う任意のコンピュータであってもよい。また、量子コンピュータと上記任意のコンピュータとが併用されて、情報処理装置10が構成されてもよい。すなわち、情報処理装置10の演算部12の全てまたは一部分が、量子コンピュータとされてもよい。

【0014】

大容量記憶装置18は、ニューラルネットワークを構成する学習モデル、後述する重要ノード特定処理を実行するためのプログラム等を記憶している。なお、大容量記憶装置18は、例えばHDD(Hard Disk Drive)や半導体記憶装置であるが、特に限定されない。

【0015】

さらに、情報処理装置10は、キーボード及びマウス等から構成され、各種操作の入力を受け付ける操作入力部、各種画像を表示する、例えば液晶ディスプレイ装置等の画像表示部、通信回線を介して他の情報処理装置等と接続され、他の情報処理装置等との間で各種データの送受信を行う外部インタフェースを備えている。

【0016】

図2は、入力情報に対する出力情報を出力する学習モデル50の一例であり、入力情報が入力される入力層52、複数のノード54で構成される中間層56、出力情報を出力する出力層58を有するニューラルネットワークとして構成される。なお、本実施形態の学

10

20

30

40

50

習モデル50は、中間層56が複数のノード54で構成されていればよく、ノード数や層数は限定されない。また、入力層52へ入力される入力情報の数や種類、出力層58から出力される出力情報等も限定されない。

【0017】

ここで、ニューラルネットワーク等の学習モデル50は、一般的にブラックボックスモデルであり、入力情報と出力情報との対応付けを説明し難い。しかしながら、入力情報と出力情報との対応付けを内部状態を元に説明できるように、学習モデル50のホワイトボックス化が望まれる場合もある。なお、ホワイトボックス化とは、入力層52と出力層58との間の中間層56を構成するノード54のうち重要なノード54を特定することで、ニューラルネットワーク等の学習モデル50の動作（入出力関係）と内部状態（特定のノード54）とを対応付けることをいう。

10

【0018】

そこで、本実施形態の情報処理装置10は、学習モデル50を構成する複数のノード54のうち除外するノード54を選択し、ノード54が除外された学習モデル50の適切度を算出する。すなわち、あるノード54が除外された学習モデル50が適切なものである場合には、除外されたノード54は相対的に重要度が低いノード54と判断される。一方、あるノード54が除外された学習モデル50が不適切なものである場合には、除外されたノード54は相対的に重要度が高いノード54と判断される。

【0019】

なお、本実施形態では学習モデル50の適切度として学習モデル50の精度（以下「モデル精度」という。）を算出する。学習モデル50のモデル精度の算出については詳細を後述する。学習モデル50の適切度とは、学習モデル50の良し悪しを評価するための指標（指標値）であり、適切度としてモデル精度以外の他の指標が用いられてもよい。

20

【0020】

ここで、ノード54の除外について図2を参照して説明する。図2において例えば中間層56における黒丸で示されるノード54が除外されるノード54（以下「除外ノード54X」という。）である。除外ノード54Xは、学習モデル50を用いた入力情報に対する出力情報の算出には寄与しない。そして、ノード54が除外された学習モデル50のモデル精度は、ノード54が除外されない学習モデル50に比べて低下する。

【0021】

ノード54の除外による学習モデル50のモデル精度の低下について具体例を挙げて説明する。例えば、学習モデル50への入力情報を、車両を運転する運転者の顔画像データとし、出力情報を運転者に眠気が生じているか否を示す情報（以下「眠気分類」という。）とする。そして、この例におけるノード54を、顔画像データに対するフィルタとする。

30

【0022】

そして、何れのノード54も除外されない学習モデル50における眠気分類の認識精度を99%とした場合、あるノード54を除外した学習モデル50における眠気分類の認識精度が例えば85%に変化したとする。また、他のノード54を除外した学習モデル50における眠気分類の認識精度が例えば95%に変化したとする。この場合、より認識精度の低下を招いたノード54が当該学習モデル50にとってより重要となるノード54（以下「重要ノード」という。）であると判断できる。このように、重要ノードとは、学習モデル50にとって出力情報に与える影響が大きいノード54であり、換言すると、学習モデル50の構造を説明することを可能とするノード54である。

40

【0023】

図3, 4を参照して、ノード54（フィルタ）を除外した学習モデル50の認識精度の試験結果を示す。なお、試験に用いた学習モデル50は、一例として、文字認識を行うものであり、図3, 4は、この学習モデル50による文字の認識精度の試験結果である。そして、図3は、一つのノード54（フィルタ）を除外した場合における学習モデル50の認識精度を示す試験結果であり、図4は、二つのノード54（フィルタ）を除外した場合における学習モデル50の認識精度を示す試験結果である。なお、図3, 4共にノード5

50

4であるフィルタの数は、30枚である。

【0024】

図3は、除外するフィルタの数が1つであるので、除外するフィルタを1つずつ選択する全探索、すなわち計30のフィルタの組み合わせにおける学習モデル50の認識精度を算出した結果である。その結果、28番目のフィルタが最も大きな認識精度の低下(約2%)を示した。なお、文字の認識精度が相対的に大きく低下したフィルタは、縦や横を示すエッジフィルタであった。

【0025】

一方、図4は、除外するフィルタの数が2つであるので、除外するフィルタを2つずつ選択する全探索、すなわち計435のフィルタの組み合わせにおける学習モデル50の認識精度を算出した結果である。その結果、文字の認識精度の低下は最大で約15%であった。この認識精度の低下は、除外するフィルタを一つとする場合に比べて2倍以上(約7倍)であり、低下の度合いがより大きくなった。この理由は、学習モデル50のモデル精度に対して一つのノード54のみが大きな影響を与えるのではなく、複数のノード54の組み合わせがモデル精度に対してより大きな影響を与えることを示している。なお、図4の試験結果でも、文字の認識精度が相対的に大きく低下したフィルタは、縦や横を示すエッジフィルタであった。

【0026】

このように、除外されることでモデル精度の低下が相対的に大きくなるノード54が、この学習モデル50にとって相対的に重要なノード54であると特定できる。従って、本実施形態の情報処理装置10は、学習モデル50のモデル精度にとって重要となるノード54を特定できるため、学習モデル50のホワイトボックス化、換言すると入力情報から出力情報へ至る理由をより明確に説明できる。また、学習モデル50から除外するノード54を複数選択し、学習モデル50のモデル精度を算出することにより、ノード54の組み合わせが学習モデル50のモデル精度に与える影響をより明確に説明できる。

【0027】

なお、以下の説明において、上述した学習モデル50における重要ノードを特定する処理を重要ノード特定処理という。

【0028】

次に図1を参照して、本実施形態の情報処理装置10が備える演算部12による重要ノード特定処理を実行する機能の詳細を説明する。演算部12は、重要ノード特定処理を実行するためにノード選択部20、モデル精度算出部24、回帰式生成部26、繰返制御部28、及び重要ノード特定部30を備える。

【0029】

なお、本実施形態の重要ノード特定処理は、学習モデル50を構成する複数のノード54の探索結果に基づいて、相対的に重要とされるノード54を特定する。すなわち、重要ノード特定処理は、除外ノード54Xを探索により一又は複数選択し、除外ノード54Xが除外された学習モデル50(以下「ノード除外学習モデル」という。)のモデル精度を算出することを繰り返し、学習モデル50にとって重要ノードを特定する。なお、ノード54の探索は、学習モデル50を構成する複数のノード54を全て探索(全探索)してもよいし、明らかに重要ではないノード54が予め特定できている場合には、当該ノード54を含まない探索(一部探索)でもよい。

【0030】

また、本実施形態の情報処理装置10は、重要ノードの特定を行うための探索として、ブラックボックス最適化を用いる。なお、ブラックボックス最適化とは、入出力関係を表す目的関数が明示的に与えられず、個別の入力に対する出力のみが与えられる場合において出力を最大化(又は最小化)する入力と出力とを求める手法である。

【0031】

この概要として、情報処理装置10は、一又は複数の除外ノード54Xを選択し、ノード除外学習モデルに評価データを適用し、ノード除外学習モデルのモデル精度を算出する

10

20

30

40

50

。そして、情報処理装置 10 は、ノード除外学習モデルを構成するノード 54 の組み合わせとノード除外学習モデルのモデル精度との関係から回帰式を生成し、この回帰式の精度（以下「予測精度」という。）を向上させるように除外ノード 54 X を選択して、ノード除外学習モデルのモデル精度を求める、という処理を繰り返し行い、回帰式をより良好なものにしていき、最終的にモデル精度がより低くなる学習モデル 50 を構成するノード 54 の組み合わせを求める。すなわち、上記処理の繰り返しは、回帰式の予測精度を上げることが目的としており、回帰式の予測精度が十分に高くなることで、モデル精度がより低くなる学習モデル 50 を構成するノード 54 の組み合わせが決定される。

【0032】

以下、演算部 12 が備える各構成部について説明する。

10

【0033】

ノード選択部 20 は、学習モデル 50 を構成する複数のノード 54 のうち、学習モデル 50 から除外するノード 54（除外ノード 54 X）を選択する。なお、選択される除外ノード 54 X の数は、一又は複数であり、以下の説明では、選択される一又は複数の除外ノード 54 X を除外ノード群ともいう。なお、1 回の選択で除外されるノード 54 の数には上限がある。この上限は、任意に設定可能とされる。

【0034】

また、ノード選択部 20 は、繰り返し処理の最初はランダムに除外ノード 54 X を選択してもよい。そして、ノード選択部 20 は、繰り返し処理の 2 回目以降は、最小化関数に基づいて除外ノード 54 X を選択するが、これについては後述する。

20

【0035】

また、本実施形態における除外ノード 54 X の選択とは、詳細を後述するようにノード 54 の選択識別子を“0”又は“1”で示すことをいうが、これに限らず、除外ノード 54 X の選択は、ノード 54 の符号の反転等、他の手法が用いられてもよい。

【0036】

モデル精度算出部 24 は、ノード選択部 20 によって選択されたノード 54 が除外された学習モデル 50（ノード除外学習モデル）に評価データを適用し、ノード除外学習モデルのモデル精度を算出する。

【0037】

なお、ノード除外学習モデルに適用する評価データは、ノード 54 が除外されていない学習モデル 50 のモデル精度の算出に用いたデータと同じデータとされる。これにより、ノード除外学習モデルのモデル精度をノード 54 が除外されていない学習モデル 50 のモデル精度と比較して判断可能となる。

30

【0038】

回帰式生成部 26 は、ノード除外学習モデルを構成するノード 54 の組み合わせ S_c とノード除外学習モデルのモデル精度との関係を表す回帰式を生成する。なお、上記関係とは、換言すると、選択された除外ノード 54 X とノード除外学習モデルのモデル精度との関係である。ここで、組み合わせ S_c を 1 回選択した段階では、除外ノード 54 X とモデル精度との関係を示すデータは一つであるが、除外ノード 54 X の選択とモデル精度の算出とを繰り返し行うことで、除外ノード 54 X とモデル精度との関係を示すデータは繰り返し数分だけ蓄積されていく。回帰式生成部 26 は、このように蓄積されたデータを用いて、除外ノード 54 X とモデル精度との関係を表す回帰式を生成する。すなわち、データが蓄積されることで回帰式の予測精度が向上する。

40

【0039】

具体的には、回帰式生成部 26 は、ノード 54 の選択識別子を s_i, s_j とした下記（1）式で示される回帰式を生成する。なお、 s_i, s_j は、ノード 54 の有無（ $\{1, 0\}$ ）を示し、ノード除外学習モデルを構成するノード 54 は“0”となり、除外ノード 54 X として選択されたノード 54 は“1”となる。

【数 1】

50

$$z = \sum_{i,j} a_{ij}s_i s_j + \sum_i b_i s_i \quad \dots (1)$$

a_{ij}, b_i : 重みパラメータ

【0040】

本構成によれば、ノード s_i, s_j と重みパラメータ a_{ij} との積の総和と、ノード s_i と重みパラメータ b_i との積の総和との和によって、ノード除外学習モデルのモデル精度を予測する回帰式が表される。なお、(1)式で示される回帰式は、モデル精度の最小化に用いられるモデル式であり、以下では最小化関数ともいう。

10

【0041】

なお、本実施形態の(1)式で示される回帰式は、2次の項で構成されるが、これに限らず回帰式は3次以上の項を含んでもよい。これによれば、2次の項で表される回帰式よりもより、より正確に重要ノードを特定できる。なお、3次以上の項として補助変数を用いることで、回帰式を2次の項で表してもよい。

【0042】

繰返制御部28は、モデル精度算出部24及び回帰式生成部26での処理を少なくとも1回以上行う。より具体的には、繰返制御部28は、回帰式生成部26によって生成された回帰式がノード除外学習モデルのモデル精度をより正確に予測するように、換言すると、回帰式によるノード除外学習モデルのモデル精度に対する予測精度が向上するようにノード選択部20によって除外ノード54Xを選択させる。そして繰返制御部28は、所定の終了条件を満たすまで、モデル精度算出部24及び回帰式生成部26での処理を繰り返し行わせる。

20

【0043】

重要ノード特定部30は、ノード除外学習モデルのモデル精度に基づいて、ノード除外学習モデルを構成する複数のノード54の中から、相対的に重要とされるノード54を特定する。なお、本実施形態の重要ノード特定部30は、ノード除外学習モデルのモデル精度の低下に基づいて、相対的に重要とされるノード54を特定する。

【0044】

なお、本実施形態の重要ノード特定部30は、所定の終了条件を満たした時点で、モデル精度がより低下したノード除外学習モデルに含まれないノード54を、重要とされるノード54として特定する。より具体的には、本実施形態の重要ノード特定部30は、一例として、モデル精度が所定値以下となったノード除外学習モデルに含まれないノード54を重要ノードとして特定する。

30

【0045】

ここで、図5は、繰り返し処理によって、ノード除外学習モデルを構成するノード54 (N_{nm}) の組み合わせ S (組み合わせ $S_{c1} \sim S_{cn}$) とモデル精度 Z (モデル精度 $z_1 \sim z_n$) のデータが蓄積されていく様子を示す模式図である。

【0046】

まず、ノード54の最初の組み合わせ S_{c1} で構成されるノード除外学習モデルに評価データを適用して求めたモデル精度は、 z_1 となる。この結果に基づいて回帰式が生成され、ノード選択部20は、回帰式に基づいてノード54の2回目の組み合わせ S_{c2} を決定し、これらのノード54で構成されるノード除外学習モデルのモデル精度は z_2 となる。3回目では、1回目の組み合わせ S_{c1} とモデル精度 z_1 と、2回目の組み合わせ S_{c2} とモデル精度 z_2 との両方を用いて回帰式が生成される。これにより求まる回帰式は、用いるデータが多い分だけ、前回の回帰式よりも予測精度が高くなる。続いて、ノード選択部20は、回帰式に基づいて3回目の組み合わせ S_{c3} を求め、選択された組み合わせ S_{c3} について、同じ処理を繰り返す。

40

【0047】

本実施形態では、複数層で構成される中間層56に対して1層毎に1又は複数の除外ノ

50

ード54Xが選択される。これにより、重要ノードを層毎に特定できる。なお、これに限らず、複数層から複数の除外ノード54Xが選択されてもよい。

【0048】

このように、組み合わせ S_{c_n} の選択、モデル精度 z_n の算出を繰り返すことにより、回帰式を生成するためのデータが増加し、組み合わせ S とモデル精度 Z との関係を示すモデル式としてより正確な回帰式が生成される。

【0049】

次に、回帰式の予測精度が高くなる、という点について説明をする。回帰式生成部26は、最小化関数である回帰式の重みパラメータ a_{ij}, b_i を図6に示されるように確率分布 $P(a_{ij}), P(b_i)$ に基づいて統計的に求める。なお、図6は、重みパラメータ a_{ij} の確率分布 $P(a_{ij})$ を示す図であるが、重みパラメータ b_i の確率分布 $P(b_i)$ も同様である。組み合わせ S_{c_n} とそれに対するモデル精度 z_n のデータが少ないと、重みパラメータ a_{ij}, b_i がとる確率分布 $P(a_{ij}), P(b_i)$ はブロードな分布であるが、除外ノード54Xの選択およびモデル精度 z の算出を繰り返し行い、データ数(図5参照)が増加するに連れて重みパラメータ a_{ij}, b_i の確率分布 $P(a_{ij}), P(b_i)$ はシャープな形状となる。すなわち、回帰式の予測精度が高くなる。

【0050】

なお、ノード選択部20は、除外ノード54Xを選択する際に、確率分布 $P(a_{ij})$ で示される重みパラメータ a_{ij}, b_i の代表値を用いて、最小化関数を最小とする除外ノード54Xを求める。一例として、最小化関数の重みパラメータ a_{ij}, b_i として適用される代表値は、確率分布 $P(a_{ij})$ に従ってサンプリングした値とする。

【0051】

また、組み合わせ S_{c_n} は回帰式の予測精度を向上させるために最小化関数によって選択するが、回帰式の予測精度を向上させる選択手段であれば、最小化関数による最小化以外の手段が用いられてもよい。一例として、効率的でないがランダムな組合せを選択することも可能である。

【0052】

次に、重要ノード特定処理について図7を用いて説明する。図7は、情報処理装置10で実行される重要ノード特定処理の流れを示すフローチャートである。また、重要ノード特定処理は、情報処理装置10が備える大容量記憶装置18等の記録媒体に格納されたプログラムによって実行される。なお、このプログラムが実行されることで、プログラムに対応する方法が実行される。

【0053】

まず、ステップS100では、ノード選択部20が、最初の除外ノード群を選択する。最初の除外ノード群は、一例として、ランダムに選択される。

【0054】

次のステップS102では、モデル精度算出部24が、除外ノード群を除外したノード除外学習モデルに対して、評価データを適用してモデル精度 z を算出する。

【0055】

次のステップS104では、回帰式生成部26が、ノード除外学習モデルを構成するノード54の組み合わせ S_{c_n} ($n=1, \dots, i$: i は繰返し数)とそのモデル精度 z_n とに基づいて、上記(1)式に示される回帰式である最小化関数を生成する。

【0056】

次のステップS106では、終了条件を満たしたか否かを繰返制御部28が判定し、肯定判定の場合はステップS110へ移行し、否定判定の場合はステップ108へ移行する。なお、終了条件は、例えば、ノード除外学習モデルを構成するノード54の組み合わせ S_{c_n} が所定数に達した場合、あるいはモデル精度 z が所定の基準以下となった場合である。

【0057】

ステップS108では、ノード選択部20が、回帰式の予測精度が向上する(本実施形

10

20

30

40

50

態ではモデル精度 z の予測値と実測値のずれが小さく、特に最小値周辺でのずれがより小さくなる) 除外ノード群を選択し、ノード除外学習モデルを構成するノード 54 の組み合わせ S_{c_n} を決定する。

【0058】

本実施形態では、ノード選択部 20 が、下記(2)式に従って、次の除外ノード群を決定する。

【数2】

$$\operatorname{argmin}_{s \in \{0,1\}^N} \sum_{i,j} a_{ij} s_i s_j + \sum_i b_i s_i + \lambda \|s\|_1 \quad \dots (2)$$

10

【0059】

(2)式は、学習モデル 50 に含まれる全てのノード 54の中から、どのノード 54 を組み合わせ S_{c_n} に含めて、どのノード 54 を組み合わせ S_{c_n} に含めないようにすれば、ノード除外学習モデルのモデル精度 z が最小となるかを求める式である。なお、第3項は正則化パラメータであり、この値を変化させることで組み合わせ S_{c_n} に含めるノード 54 の数を調整することができる。すなわち、正則化パラメータの値を調整することにより、除外ノード 54 X の数を調整することができる。なお、除外ノード 54 X の数は、必ずしも正則化パラメータによって定義する必要は無く、予め設定されてもよい。

【0060】

20

ここで、繰り返し回数が少ないときには回帰式の予測精度は低く、除外ノード群を除外した組み合わせ S_{c_i} は、ランダムに選択した除外ノード群を除外した場合に相当する。一方で繰り返し回数が増えると回帰式の予測精度は高くなり、組み合わせ S_{c_i} は、最終的に求めたいノード除外学習モデルに相当する。この特性により、可能な組み合わせ S_{c_i} を広くカバーし、かつ求めたい組み合わせ S_{c_i} の周辺を高い精度で予測できる回帰式を効率的に得ることができる。なお、この手法に限らず、効率的に高い予測精度の回帰式が得られれば他の手段が用いられてもよい。

【0061】

そして、重要ノード特定処理は、新しい除外ノード群を選択したら、その除外ノード群を除外したノード 54 で構成されるノード除外学習モデルのモデル精度 z の算出(S102)、回帰式の生成(S104)を行い、終了条件を満たすまで、これらの処理を繰り返す。

30

【0062】

一方、ステップ S106における判定で終了条件を満たした場合、ステップ S110において重要ノード特定部 30 が、終了条件を満たした時点におけるモデル精度 z が所定値以下となるノード除外学習モデルに含まれないノード 54 を重要ノードとして特定する。

【0063】

なお、本実施形態では、上述のように、除外ノード 54 X を複数選択することも実行する。これにより、重要ノード特定処理は、学習モデル 50 のモデル精度 z に影響を与える一つのノード 54 を特定するのではなく、ノード 54 の組み合わせで特定することが可能となる。これにより、学習モデル 50 のホワイトボックス化、換言すると入力情報から出力情報へ至る理由をより明確に説明できる。

40

【0064】

以上、本発明を、上記実施形態を用いて説明したが、本発明の技術的範囲は上記実施形態に記載の範囲には限定されない。発明の要旨を逸脱しない範囲で上記実施形態に多様な変更又は改良を加えることができ、該変更又は改良を加えた形態も本発明の技術的範囲に含まれる。

【0065】

上記実施形態では、探索条件として(2)式に示されるように“argmin”を適用する形態について説明したが、本発明はこれに限られず、探索条件として“argmax”を適用してもよ

50

い。例えば、ノード除外学習モデルのモデル精度 z とノード 54 を除外していない学習モデル 50 のモデル精度（基準値）との差が大きくなるように除外ノード群を選択してもよく、このような場合に、探索条件として “argmax” が適用される。

【0066】

また、上記実施形態では、ニューラルネットワークの中間層 56 における重要ノードを特定する形態について説明したが、本発明はこれに限られない。例えば、ニューラルネットワークの入力層 52 が複数のノード 54 で構成される場合には、入力層 52 における重要ノードを特定する形態としてもよい。

【0067】

また、上記実施形態では、入力情報を車両の運転者の顔画像データとし、出力情報を運転者の眠気分類とする形態について説明したが、本発明はこれに限られず、入力情報や出力情報は他の情報であってもよく、限定されるものではない。

10

【0068】

また、上記実施形態では、回帰モデルとして（1）式に示される確率分布を用いた回帰式を適用する形態について説明したが、本発明はこれに限られない。例えば、回帰モデルとして Factorization Machine 等の他の手法（アルゴリズム）を適用してもよい。

【0069】

また、上記実施形態では、重要ノードと特定するためにブラックボックス最適化を適用する形態について説明したが、本発明はこれに限られず、重要ノードと特定することができる演算処理であれば、他の処理方法が用いられてもよい。

20

【0070】

また、上記実施形態では、学習モデル 50 の適切度としてモデル精度を適用する形態について説明したが、本発明はこれに限られず、学習モデル 50 の適切度として数値で表すことができる他の指標（指標値）が用いられてもよい。例えば、学習モデル 50 が入力データの分類を行う場合には、本実施形態のモデル精度に相当する正解率の他に、適合率、再現率、F 値、混同行列、マイクロ平均値、マクロ平均値、ROC (Receiver Operating Characteristic) 曲線、AUC (Area Under the Curve) が適用されてもよい。また、この他にも学習の評価指標として loss が適用されたり、アプリケーションの評価指標として mIoU (mean IoU) や画像比較の結果等が適用されてもよい。

【0071】

30

また、上記実施形態では、ノード除外学習モデルのモデル精度の低下に基づいて、相対的に重要とされるノード 54 を特定する形態について説明したが、本発明はこれに限られない。例えば、重要ノードは、ノード除外学習モデルのモデル精度の変化に基づいて特定されればよく、例えば、ノード除外学習モデルのモデル精度の上昇に基づいて、相対的に重要とされるノード 54 が特定されてもよい。

【産業上の利用可能性】

【0072】

本発明は、ニューラルネットワークのように複数のノードで構成される学習モデルに用いることができる。

【符号の説明】

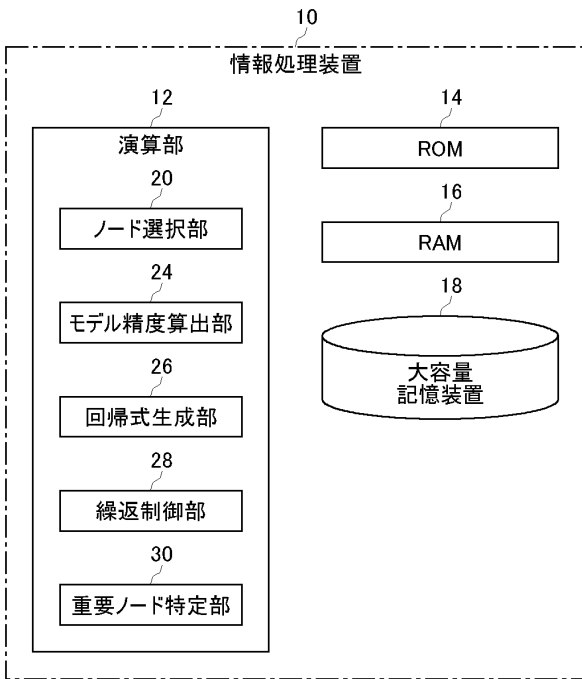
40

【0073】

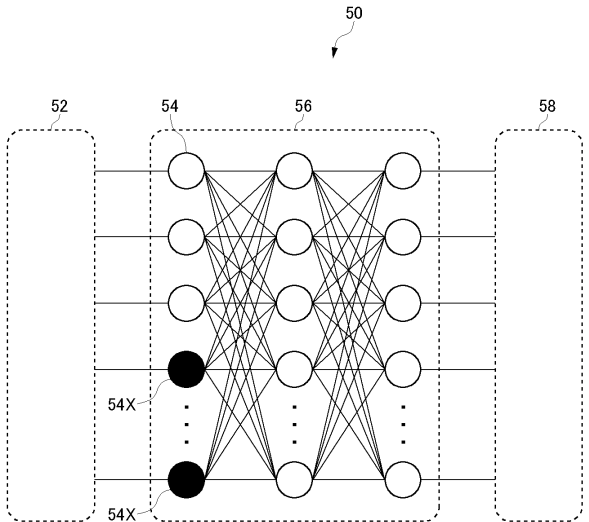
10・・・情報処理装置、20・・・選択部、24・・・モデル精度算出部、
26・・・回帰式生成部、28・・・繰返制御部、30・・・重要ノード特定部、
50・・・学習モデル、54・・・ノード、54X・・・除外ノード

【図面】

【図 1】



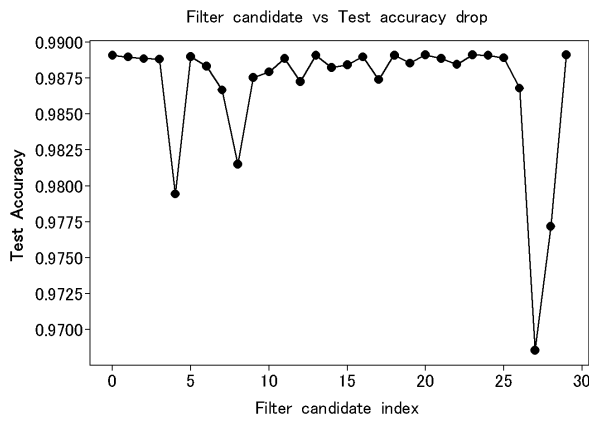
【図 2】



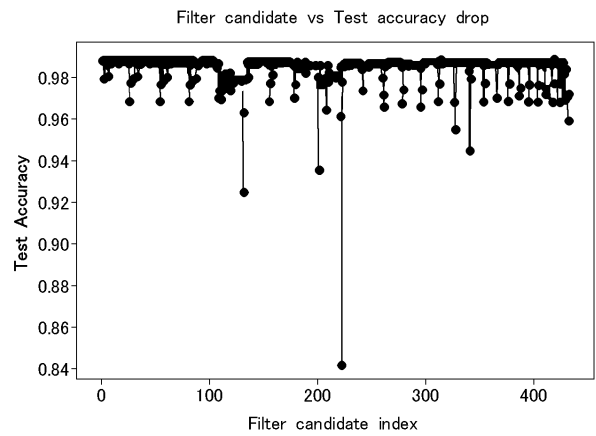
10

20

【図 3】



【図 4】



30

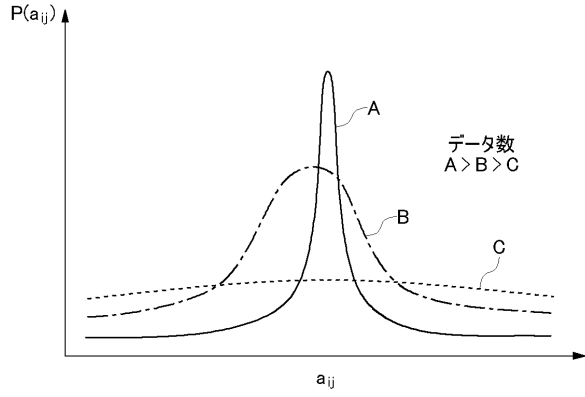
40

50

【図5】

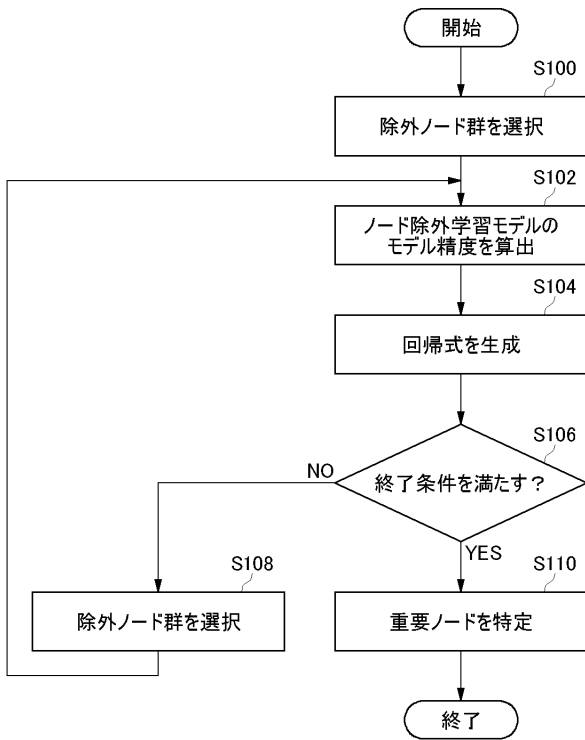
	N_{11}	N_{12}	N_{13}	...	N_{21}	N_{22}	N_{23}	...	N_{31}	N_{32}	N_{33}	...
Sc_1	1	0	1	...	0	0	0	...	0	0	0	...
Sc_2	0	1	0	...	0	0	0	...	0	0	0	...
Sc_3	1	1	0	...	0	0	0	...	0	0	0	...
⋮	⋮											
	S											Z

【図6】



10

【図7】



20

30

40

50

フロントページの続き

審査官 武田 広太郎

- (56)参考文献 特開平09-091263(JP,A)
国際公開第2017/131081(WO,A1)
- (58)調査した分野 (Int.Cl., DB名)
- | | |
|------|-------|
| G06N | 3/082 |
| G06N | 3/08 |