

(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号
特許第7344900号
(P7344900)

(45)発行日 令和5年9月14日(2023.9.14)

(24)登録日 令和5年9月6日(2023.9.6)

(51)国際特許分類 F I
G 0 6 N 3/045(2023.01) G 0 6 N 3/045

請求項の数 20 (全25頁)

(21)出願番号	特願2020-555022(P2020-555022)	(73)特許権者	314015767
(86)(22)出願日	平成31年4月27日(2019.4.27)		マイクロソフト テクノロジー ライセン
(65)公表番号	特表2021-523430(P2021-523430 A)		シング,エルエルシー
(43)公表日	令和3年9月2日(2021.9.2)		アメリカ合衆国 ワシントン州 9 8 0 5
(86)国際出願番号	PCT/US2019/029532		2 レッドモンド ワン マイクロソフト
(87)国際公開番号	WO2019/217113	(74)代理人	100079108
(87)国際公開日	令和1年11月14日(2019.11.14)		弁理士 稲葉 良幸
審査請求日	令和4年3月28日(2022.3.28)	(74)代理人	100109346
(31)優先権主張番号	15/976,514		弁理士 大貫 敏史
(32)優先日	平成30年5月10日(2018.5.10)	(74)代理人	100117189
(33)優先権主張国・地域又は機関	米国(US)		弁理士 江口 昭彦
		(74)代理人	100134120
			弁理士 内藤 和彦
		(74)代理人	100108213

最終頁に続く

(54)【発明の名称】 教師付き機械学習問題用のニューラルネットワークアーキテクチャの選択

(57)【特許請求の範囲】

【請求項 1】

処理ハードウェア、並びに

機械学習問題と関連付けられた機械学習問題空間、及び前記機械学習問題を解くための複数の訓練を受けていない候補ニューラルネットワークにアクセスすること、

訓練を受けていない候補ニューラルネットワークごとに、前記機械学習問題に対する前記候補ニューラルネットワークの表現性を捕捉する少なくとも1つの表現性測度を計算すること、

訓練を受けていない候補ニューラルネットワークごとに、前記機械学習問題に対する前記候補ニューラルネットワークの訓練可能性を捕捉する少なくとも1つの訓練可能性測度を計算すること、

前記少なくとも1つの表現性測度及び前記少なくとも1つの訓練可能性測度に基づいて、前記機械学習問題を解くための少なくとも1つの候補ニューラルネットワークを選択すること、

前記選択された少なくとも1つの候補ニューラルネットワークを表す出力を提供すること、

を含む動作を前記処理ハードウェアに行わせる命令を保存するメモリを含む、システムであって、前記少なくとも1つの訓練可能性測度は、勾配変形を表す、システム。

【請求項 2】

処理ハードウェア、並びに

10

20

機械学習問題と関連付けられた機械学習問題空間、及び前記機械学習問題を解くための複数の訓練を受けていない候補ニューラルネットワークにアクセスすることと、
訓練を受けていない候補ニューラルネットワークごとに、前記機械学習問題に対する前記候補ニューラルネットワークの表現性を捕捉する少なくとも1つの表現性測度を計算することと、

訓練を受けていない候補ニューラルネットワークごとに、前記機械学習問題に対する前記候補ニューラルネットワークの訓練可能性を捕捉する少なくとも1つの訓練可能性測度を計算することと、

前記少なくとも1つの表現性測度及び前記少なくとも1つの訓練可能性測度に基づいて、前記機械学習問題を解くための少なくとも1つの候補ニューラルネットワークを選択することと、

10

前記選択された少なくとも1つの候補ニューラルネットワークを表す出力を提供することと、

を含む動作を前記処理ハードウェアに行わせる命令を保存するメモリを含む、システムであって、前記少なくとも1つの訓練可能性測度が、訓練の第1段階の間、前記候補ニューラルネットワークにおける重みの確率的勾配降下を表す、システム。

【請求項3】

前記少なくとも1つの表現性測度が、前記訓練を受けていない候補ニューラルネットワークによる、前記機械学習問題空間からのサンプルの分離の測度を表す、請求項2に記載のシステム。

20

【請求項4】

前記分離の測度が大きさである、請求項3に記載のシステム。

【請求項5】

前記分離の測度が角度である、請求項3に記載のシステム。

【請求項6】

前記機械学習問題を解くための前記少なくとも1つの候補ニューラルネットワークを選択することが、

閾値を超える前記少なくとも1つの表現性測度、及び範囲内の前記少なくとも1つの訓練可能性測度を有する前記少なくとも1つの候補ニューラルネットワークを選択することを含み、前記範囲が、範囲最小値及び範囲最大値によって定義される、請求項2に記載のシステム。

30

【請求項7】

前記動作が、

前記機械学習問題を解くために前記少なくとも1つの候補ニューラルネットワークを訓練することをさらに含む、請求項2に記載のシステム。

【請求項8】

前記動作が、

前記機械学習問題を解くために、前記訓練された少なくとも1つの候補ニューラルネットワークを前記機械学習問題空間で実行することと、

前記訓練された少なくとも1つの候補ニューラルネットワークによって生成された前記機械学習問題に対する解を提供することと、

40

をさらに含む、請求項7に記載のシステム。

【請求項9】

機械学習問題と関連付けられた機械学習問題空間、及び前記機械学習問題を解くための複数の訓練を受けていない候補ニューラルネットワークにアクセスすることと、

訓練を受けていない候補ニューラルネットワークごとに、前記機械学習問題に対する前記候補ニューラルネットワークの表現性を捕捉する少なくとも1つの表現性測度を計算することと、

訓練を受けていない候補ニューラルネットワークごとに、前記機械学習問題に対する前記候補ニューラルネットワークの訓練可能性を捕捉する少なくとも1つの訓練可能性測度

50

を計算することと、

前記少なくとも1つの表現性測度及び前記少なくとも1つの訓練可能性測度に基づいて、前記機械学習問題を解くための少なくとも1つの候補ニューラルネットワークを選択することと、

前記選択された少なくとも1つの候補ニューラルネットワークを表す出力を提供することと、

を含む動作を1つ又は複数の機械に行わせる命令を保存する、機械可読媒体であって、前記少なくとも1つの訓練可能性測度は、勾配変形を表す、機械可読媒体。

【請求項10】

機械学習問題と関連付けられた機械学習問題空間、及び前記機械学習問題を解くための複数の訓練を受けていない候補ニューラルネットワークにアクセスすることと、

訓練を受けていない候補ニューラルネットワークごとに、前記機械学習問題に対する前記候補ニューラルネットワークの表現性を捕捉する少なくとも1つの表現性測度を計算することと、

訓練を受けていない候補ニューラルネットワークごとに、前記機械学習問題に対する前記候補ニューラルネットワークの訓練可能性を捕捉する少なくとも1つの訓練可能性測度を計算することと、

前記少なくとも1つの表現性測度及び前記少なくとも1つの訓練可能性測度に基づいて、前記機械学習問題を解くための少なくとも1つの候補ニューラルネットワークを選択することと、

前記選択された少なくとも1つの候補ニューラルネットワークを表す出力を提供することと、

を含む動作を1つ又は複数の機械に行わせる命令を保存する、機械可読媒体であって、前記少なくとも1つの訓練可能性測度が、訓練の第1段階の間、前記候補ニューラルネットワークにおける重みの確率的勾配降下を表す、機械可読媒体。

【請求項11】

前記少なくとも1つの表現性測度が、前記訓練を受けていない候補ニューラルネットワークによる、前記機械学習問題空間からのサンプルの分離の測度を表す、請求項10に記載の機械可読媒体。

【請求項12】

前記分離の測度が大きさである、請求項11に記載の機械可読媒体。

【請求項13】

前記分離の測度が角度である、請求項11に記載の機械可読媒体。

【請求項14】

前記機械学習問題を解くための前記少なくとも1つの候補ニューラルネットワークを選択することが、

閾値を超える前記少なくとも1つの表現性測度、及び範囲内の前記少なくとも1つの訓練可能性測度を有する前記少なくとも1つの候補ニューラルネットワークを選択することを含み、前記範囲が、範囲最小値及び範囲最大値によって定義される、請求項10に記載の機械可読媒体。

【請求項15】

処理ハードウェアによって実行される方法であって、

機械学習問題と関連付けられた機械学習問題空間、及び前記機械学習問題を解くための複数の訓練を受けていない候補ニューラルネットワークにアクセスすることと、

訓練を受けていない候補ニューラルネットワークごとに、前記機械学習問題に対する前記候補ニューラルネットワークの表現性を捕捉する少なくとも1つの表現性測度を計算することと、

訓練を受けていない候補ニューラルネットワークごとに、前記機械学習問題に対する前記候補ニューラルネットワークの訓練可能性を捕捉する少なくとも1つの訓練可能性測度を計算することと、

10

20

30

40

50

前記少なくとも1つの表現性測度及び前記少なくとも1つの訓練可能性測度に基づいて、前記機械学習問題を解くための少なくとも1つの候補ニューラルネットワークを選択することと、

前記選択された少なくとも1つの候補ニューラルネットワークを表す出力を提供することと、

を含み、前記少なくとも1つの訓練可能性測度は、勾配変形を表す、方法。

【請求項16】

処理ハードウェアによって実行される方法であって、

機械学習問題と関連付けられた機械学習問題空間、及び前記機械学習問題を解くための複数の訓練を受けていない候補ニューラルネットワークにアクセスすることと、

訓練を受けていない候補ニューラルネットワークごとに、前記機械学習問題に対する前記候補ニューラルネットワークの表現性を捕捉する少なくとも1つの表現性測度を計算することと、

訓練を受けていない候補ニューラルネットワークごとに、前記機械学習問題に対する前記候補ニューラルネットワークの訓練可能性を捕捉する少なくとも1つの訓練可能性測度を計算することと、

前記少なくとも1つの表現性測度及び前記少なくとも1つの訓練可能性測度に基づいて、前記機械学習問題を解くための少なくとも1つの候補ニューラルネットワークを選択することと、

前記選択された少なくとも1つの候補ニューラルネットワークを表す出力を提供することと、

を含み、前記少なくとも1つの訓練可能性測度が、訓練の第1段階の間、前記候補ニューラルネットワークにおける重みの確率的勾配降下を表す、方法。

【請求項17】

前記少なくとも1つの表現性測度が、前記訓練を受けていない候補ニューラルネットワークによる、前記機械学習問題空間からのサンプルの分離の測度を表す、請求項16に記載の方法。

【請求項18】

前記分離の測度が大きさである、請求項17に記載の方法。

【請求項19】

前記分離の測度が角度である、請求項17に記載の方法。

【請求項20】

前記機械学習問題を解くための前記少なくとも1つの候補ニューラルネットワークを選択することが、

閾値を超える前記少なくとも1つの表現性測度、及び範囲内の前記少なくとも1つの訓練可能性測度を有する前記少なくとも1つの候補ニューラルネットワークを選択することを含み、前記範囲が、範囲最小値及び範囲最大値によって定義される、請求項16に記載の方法。

【発明の詳細な説明】

【背景技術】

【0001】

背景

【0001】 複数の異なるタイプのニューラルネットワークアーキテクチャ（例えば、畳み込みニューラルネットワーク、フィードフォワードニューラルネットワークなど）が知られている。所与の機械学習問題を解くためのニューラルネットワークアーキテクチャ（及び所与のアーキテクチャタイプ内のサブアーキテクチャ）を選択することは、困難なことがある。

【0002】

図面の簡単な説明

【0002】 本技術の幾つかの実施形態は、例として、限定ではなく添付の図面の図に示さ

10

20

30

40

50

れる。

【図面の簡単な説明】

【0003】

【図1】[0003]幾つかの実施形態による、機械学習問題を解くためのニューラルネットワークアーキテクチャを選択することが実施され得る例示的システムを示す。

【図2】[0004]幾つかの実施形態による、機械学習問題を解くためのニューラルネットワークアーキテクチャを選択する例示的方法のフローチャートを示す。

【図3】[0005]幾つかの実施形態による、誤り率を減少させる例示的方法のフローチャートを示す。

【図4A】[0006]幾つかの実施形態による、ブロックの中からセルを選択する例示的方法のフローチャートを示す。

10

【図4B】[0006]幾つかの実施形態による、ブロックの中からセルを選択する例示的方法のフローチャートを示す。

【図5】[0007]幾つかの実施形態による、機械可読媒体から命令を読み出し、本明細書で述べる方法論の何れかを行うことが可能な機械のコンポーネントを示すブロック図である。

【発明の概要】

【発明が解決しようとする課題】

【0004】

概要

[0008] 本開示は、一般に、機械学習問題を解くためのニューラルネットワークアーキテクチャを選択するように構成された機械（このような専用機械のコンピュータ化された変形形態及びそのような変形形態の改良形態を含む）、並びにそのような専用機械がニューラルネットワークの技術を提供する他の専用機械と比較して改良される技術に関する。具体的には、本開示は、所与の機械学習問題空間において機械学習問題を解くためのニューラルネットワークアーキテクチャを選択するシステム及び方法に取り組む。

20

【課題を解決するための手段】

【0005】

[0009] 本明細書に記載する技術の幾つかの局面によれば、システムは、処理ハードウェア及びメモリを含む。メモリは、処理ハードウェアによって実行されると、処理ハードウェアに動作を行わせる命令を保存する。動作は、機械学習問題と関連付けられた機械学習問題空間、及び機械学習問題を解くための複数の訓練を受けていない候補ニューラルネットワークにアクセスすることを含む。動作は、訓練を受けていない候補ニューラルネットワークごとに、機械学習問題に対する候補ニューラルネットワークの表現性を捕捉する少なくとも1つの表現性測度を計算することを含む。動作は、訓練を受けていない候補ニューラルネットワークごとに、機械学習問題に対する候補ニューラルネットワークの訓練可能性に関連する少なくとも1つの訓練可能性測度を計算することを含む。動作は、少なくとも1つの表現性測度及び少なくとも1つの訓練可能性測度に基づいて、機械学習問題を解くための少なくとも1つの候補ニューラルネットワークを選択することを含む。動作は、選択された少なくとも1つの候補ニューラルネットワークを表す出力を提供することを含む。

30

40

【0006】

[0010] 本明細書に記載する技術の幾つかの局面によれば、機械可読媒体は、1つ又は複数の機械によって実行されると、1つ又は複数の機械に動作を行わせる命令を保存する。動作は、機械学習問題と関連付けられた機械学習問題空間、及び機械学習問題を解くための複数の訓練を受けていない候補ニューラルネットワークにアクセスすることを含む。動作は、訓練を受けていない候補ニューラルネットワークごとに、機械学習問題に対する候補ニューラルネットワークの表現性を捕捉する少なくとも1つの表現性測度を計算することを含む。動作は、訓練を受けていない候補ニューラルネットワークごとに、機械学習問題に対する候補ニューラルネットワークの訓練可能性を捕捉する少なくとも1つの訓練可能性測度を計算することを含む。動作は、少なくとも1つの表現性測度及び少なくとも

50

1つの訓練可能性測度に基づいて、機械学習問題を解くための少なくとも1つの候補ニューラルネットワークを選択することを含む。動作は、選択された少なくとも1つの候補ニューラルネットワークを表す出力を提供することを含む。

【0007】

[0011] 本明細書に記載する技術の幾つかの局面によれば、方法は、入力行列にアクセスすることを含む。方法は、機械学習問題と関連付けられた機械学習問題空間、及び機械学習問題を解くための複数の訓練を受けていない候補ニューラルネットワークにアクセスすることを含む。方法は、訓練を受けていない候補ニューラルネットワークごとに、機械学習問題に対する候補ニューラルネットワークの表現性を捕捉する少なくとも1つの表現性測度を計算することを含む。方法は、訓練を受けていない候補ニューラルネットワークごとに、機械学習問題に対する候補ニューラルネットワークの訓練可能性を捕捉する少なくとも1つの訓練可能性測度を計算することを含む。方法は、少なくとも1つの表現性測度及び少なくとも1つの訓練可能性測度に基づいて、機械学習問題を解くための少なくとも1つの候補ニューラルネットワークを選択することを含む。方法は、選択された少なくとも1つの候補ニューラルネットワークを表す出力を提供することを含む。

10

【発明を実施するための形態】

【0008】

詳細な説明
概要

[0012] 本開示は、特に、様々な機能性を個々に提供する方法、システム、及びコンピュータプログラム製品を記載する。以下の記載では、説明目的で、本開示の異なる実施形態の様々な局面の十分な理解を提供するために、多数の具体的な詳細を記載する。しかし、これらの具体的な詳細が全てなくても本開示を実施し得ることが、当業者には明白となるだろう。

20

【0009】

[0013] 上記の通り、複数の異なるタイプのニューラルネットワークアーキテクチャ（例えば、フィードフォワードニューラルネットワーク、畳み込みニューラルネットワーク、リカレントニューラルネットワークなど）が知られている。所与の機械学習問題（例えば、回帰問題、順位付け問題、又は所与の空間内でデータを分類することなど（鳥の画像を画像中の鳥の種類によって分類することなど）の分類問題）を解くための高性能ニューラルネットワークアーキテクチャを選択することは、困難なことがある。

30

【0010】

[0014] 本明細書に記載する技術の幾つかの局面は、1組のニューラルネットワークアーキテクチャから、所与の機械学習問題を解くためのニューラルネットワークアーキテクチャをニューラルネットワークアーキテクチャが訓練される前に選択する技術的問題を解決することに向けられる。有利に、幾つかの局面の結果、所与の機械学習問題を解くために高性能ニューラルネットワークアーキテクチャが訓練され、あまり望ましくないアーキテクチャは訓練されない。このことは、非高性能ニューラルネットワークが使用されることをもたらすことなく、計算時間を節約し、効率を高める。

【0011】

40

[0015] 場合によっては、この問題に対する解決策は、サーバで実施される。サーバは、データリポジトリを介して、機械学習問題に関連付けられた機械学習問題空間、及び機械学習問題を解くための複数の訓練を受けていない候補ニューラルネットワークにアクセスする。サーバは、訓練を受けていない候補ニューラルネットワークごとに、機械学習問題に対する候補ニューラルネットワークの表現性を捕捉する少なくとも1つの表現性測度を計算する。サーバは、訓練を受けていない候補ニューラルネットワークごとに、機械学習問題に対する候補ニューラルネットワークの訓練可能性を捕捉する少なくとも1つの訓練可能性測度を計算する。サーバは、少なくとも1つの表現性測度、少なくとも1つの訓練可能性測度、及び候補ニューラルネットワークのアーキテクチャに基づいて、機械学習問題を解くための少なくとも1つの候補ニューラルネットワークを選択する。サーバは、

50

選択された少なくとも1つの候補ニューラルネットワークを表す出力を提供する。

【0012】

[0016] 場合によっては、選択された少なくとも1つの候補ニューラルネットワークは、機械学習問題を解くために部分的又は完全に訓練される。本明細書では、ニューラルネットワークが「部分的又は完全に訓練される」とは、数エポックの間、訓練される、又はある収束の指標が達成されるまで訓練されることを含み得る。訓練された少なくとも1つの候補ニューラルネットワークは、機械学習問題を解くために、機械学習問題空間に対して実行される。次いで、サーバは、訓練された少なくとも1つの候補ニューラルネットワークによって生成された機械学習問題に対する解を提供する。

【0013】

[0017] 幾つかの例によれば、少なくとも1つの表現性測度は、訓練を受けていない候補ニューラルネットワークによる、分類問題空間からのサンプルの分離の測度（例えば、大きさ又は角度）を表す。幾つかの例によれば、少なくとも1つの訓練可能性測度は、機械学習問題空間からのサンプルを所与として、最後の層及び最初の層における勾配の関数を表す。幾つかの例によれば、表現性測度及び訓練可能性測度は、表現性及び訓練可能性などのアーキテクチャの異なる特性を捕捉するニューラルネットワークの量、測度、又は統計値を含み得る。

【0014】

[0018] 図1は、幾つかの実施形態による、機械学習問題を解くためのニューラルネットワークアーキテクチャを選択することが実施され得る例示的システム100を示す。図示のように、システム100は、ネットワーク140を介して互いに接続されたサーバ110、データリポジトリ120、及びクライアントデバイス130を含む。ネットワーク140は、インターネット、イントラネット、ローカルエリアネットワーク、広域ネットワーク、有線ネットワーク、無線ネットワーク、セルラーネットワーク、WiFiネットワークなどの1つ又は複数を含む。

【0015】

[0019] クライアントデバイス130は、ラップトップコンピュータ、デスクトップコンピュータ、携帯電話、タブレットコンピュータ、プロセッサ及びメモリを備えたスマートテレビ、スマートウォッチなどでもよい。クライアントデバイス130は、出力をユーザに表示するため、又は入力をユーザから受信するために使用され得る。

【0016】

[0020] データリポジトリ120は、データベース又はその他のデータストレージ構造として実施され得る。図示のように、データリポジトリは、機械学習問題空間122を保存する。機械学習問題空間122は、ニューラルネットワークによって分類されるべきデータを含む。例えば、機械学習問題空間122は、鳥の種類によって分類されるべき鳥の写真、又は「重要な電子メール」、「重要でない電子メール」、又は「スパム」として分類されるべき電子メールメッセージを含み得る。

【0017】

[0021] サーバ110は、1つ又は複数のサーバを含み得る。サーバ110は、複数のサーバを含むサーバファームとして実施されてもよい。図示のように、サーバ110は、訓練を受けていない候補ニューラルネットワーク112.1-n（nは、2以上の正の整数である）、選択モジュール114、及び訓練モジュール116を保存する。訓練を受けていない候補ニューラルネットワーク112.1-nは、様々な分類タスクに使用され得るニューラルネットワークである。例えば、訓練を受けていない候補ニューラルネットワーク112.1-nは、畳み込みニューラルネットワーク又はフィードフォワードニューラルネットワークの訓練を受けていないバージョンを含み得る。

【0018】

[0022] 選択モジュール114は、機械学習問題空間122と関連付けられた機械学習問題を解くために訓練する、訓練を受けていない候補ニューラルネットワーク112.1-nの少なくとも1つを選択する。選択モジュールの例示的動作のさらなる詳細を図2に

10

20

30

40

50

関連して提供する。訓練モジュール 1 1 6 は、機械学習問題を解くために、(訓練を受けていない候補ニューラルネットワーク 1 1 2 . 1 - n から) (選択モジュール 1 1 4 によって) 選択された 1 つ又は複数のニューラルネットワークを訓練する。訓練後に、訓練されたニューラルネットワークを使用して、機械学習問題空間 1 2 2 (又は別の問題空間) におけるデータを分類することによって機械学習問題を解くことができる。

【 0 0 1 9 】

[0023] 図 2 は、幾つかの実施形態による、機械学習問題を解くためのニューラルネットワークアーキテクチャを選択する例示的方法 2 0 0 のフローチャートを示す。以下に記載するように、方法 2 0 0 は、図 1 のサーバ 1 1 0 の選択モジュール 1 1 4 を用いて実施される。しかし、方法 2 0 0 は、システム 1 0 0 のアーキテクチャに制限されず、他のアーキテクチャ又は他のシステムにおいて実施され得る。

10

【 0 0 2 0 】

[0024] 動作 2 1 0 では、選択モジュール 1 1 4 は、解かれるべき機械学習問題に関連付けられた機械学習問題空間 1 2 2 に (例えばネットワーク 1 4 0 を介して) アクセスする。選択モジュール 1 1 4 は、機械学習問題を解くための複数の訓練を受けていない候補ニューラルネットワーク 1 1 2 . 1 - n にアクセスする。

【 0 0 2 1 】

[0025] 動作 2 2 0 では、選択モジュール 1 1 4 は、訓練を受けていない候補ニューラルネットワーク 1 1 2 . 1 - n ごとに、機械学習問題に対する候補ニューラルネットワークの表現性に関連する 1 つ又は複数の表現性メトリックを計算する。1 つ又は複数の表現性メトリックは、訓練を受けていない候補ニューラルネットワークによる、機械学習問題空間 1 2 2 からのサンプルの分離の測度を表す。分離の測度は、大きさ又は角度でもよい。

20

【 0 0 2 2 】

[0026] 動作 2 3 0 では、選択モジュール 1 1 4 は、訓練を受けていない候補ニューラルネットワーク 1 1 2 . 1 - n ごとに、機械学習問題に対する候補ニューラルネットワークの訓練可能性に関連する 1 つ又は複数の訓練可能性メトリックを計算する。

【 0 0 2 3 】

[0027] 動作 2 4 0 では、選択モジュール 1 1 4 は、1 つ又は複数の表現性測度及び 1 つ又は複数の訓練可能性測度に基づいて、機械学習問題を解くための 1 つ又は複数の候補ニューラルネットワークを選択する。この選択は、1 つ又は複数の候補ニューラルネットワークの 1 つ又は複数のアーキテクチャにも基づく。1 つ又は複数の候補ニューラルネットワークは、複数の訓練を受けていない候補ニューラルネットワーク 1 1 2 . 1 - n から選択される。場合によっては、選択モジュール 1 1 4 は、閾値を超える 1 つ又は複数の表現性測度、及び範囲内の 1 つ又は複数の訓練可能性測度を有する、1 つ又は複数の候補ニューラルネットワークを選択する。この範囲は、範囲最小値及び範囲最大値によって定義される。

30

【 0 0 2 4 】

[0028] 動作 2 5 0 では、選択モジュール 1 1 4 は、1 つ又は複数の選択された候補ニューラルネットワークを表す出力を提供する。場合によっては、訓練モジュール 1 1 6 は、機械学習問題を解くために、1 つ又は複数の選択された候補ニューラルネットワークを訓練する。訓練モジュール 1 1 6 は、機械学習問題を解くために、機械学習問題空間 1 2 2 に対して 1 つ又は複数の訓練された候補ニューラルネットワークを実行する。サーバ 1 1 0 は、1 つ又は複数の訓練された候補ニューラルネットワークによって生成された機械学習問題に対する解を (例えば、クライアントデバイス 1 3 0 で表示するためにクライアントデバイス 1 3 0 に対して、又はデータリポジトリ 1 2 0 で保存するためにデータリポジトリ 1 2 0 に対して) 提供する。

40

【 0 0 2 5 】

[0029] 本明細書に記載する技術の目標の 1 つは、所与のタスク用のニューラルネットワークアーキテクチャの選択及び構成を自動的に行うことである。より具体的には、データセットを所与として、幾つかの局面は、層のタイプ (例えば、畳み込み、マックスプー

50

リング、完全接続など)、それらのハイパーパラメータ(例えば、ストライドサイズ、畳み込みサイズ)、他の全ての層に対するそれらの接続、及び層の総数を自動的に識別する。幾つかの局面は、どの訓練アルゴリズムを使用するか(例えば、確率的勾配降下、RMS Prop、Adamなど)、及び重みの初期化方法(例えば、Glorot、Normal、Laplace、Uniformなど)も識別する。

【0026】

[0030] 幾つかのスキームは、(1)それらが複数のアーキテクチャにわたりサーチ空間をどのように定義するか(例えば、無制限又は制限)、(2)それらが空間をどのように探索するか(例えば、強化学習、モンテカルロ木探索)、(3)それらがサーチを誘導するためにどの予測モデルを使用するか(例えば、sequential model based optimization、リカレントニューラルネットワーク(RNN)、遺伝的アルゴリズム)、及び(4)それらがサーチをより効率的に誘導するために安価な代理関数を使用するか否か、に基づいて、グループ化され得る。遺伝的アルゴリズム及び強化学習の場合、上記の項目2及び3は、それらが一緒に予測モデルを学習し、空間を探索するため、1つにまとめられる。

10

【0027】

[0031] 幾つかの実施形態では、サーチ空間が、sequential model-based optimization(SMBO)を用いて探索される。SMBOの例示的モデルは、ベイジアンリカレントニューラルネットワークでもよい。1組のランダムニューラルネットワーク特性が、ニューラルネットワークの真の性能に対する安価なサロゲート(又は統計値)として機能すると定義される。場合によっては、予測及び予測に関する不確実性の両方を出力する他のモデル(例えば、ガウス過程)が使用される。場合によっては、安価なサロゲート(又は統計値)が、モデルの入力として与えられるのではなく、出力と見なされる。安価なサロゲートは、上述の訓練可能性測度及び表現性測度を含み得る。

20

【0028】

[0032] 幾つかの拡張実施形態では、サーチ空間は、無制限である。強化学習又はモンテカルロ木探索を用いて、サーチ空間が探索される。場合によっては、予測を支援するためにアーキテクチャ自体の表現を追加的に使用することも可能である。

【0029】

[0033] 幾つかの局面は、訓練前に、訓練後のデータセットに関する初期モデルの性能を予測するための2つの主要な測度に向けられる。両測度は、データセットからのランダムバッチに関して収集されたモデルの統計値である。幾つかの例では、バッチは、n個の点から成る

30

【数1】

$$\{(x_i, y_i)\}_{i=1}^n$$

(幾つかの局面は、ラベル y_i を無視する)。モデルは、従来、2つのコンポーネント：入力空間を潜在空間に埋め込むディープニューラルネットワーク、及び埋め込みを1組の可能なラベルに関する確率分布に変化させるソフトマックスが後に続く完全接続線形層、を含む。fを前者(すなわち、埋め込み)とし、それがL個の層を有すると仮定する。

40

【0030】

[0034] メトリック表現性は、式1(これは、式2に近似する)に従って定義される。式1では、 $\{x_1, \dots, x_n\}$ は、入力のバッチである。

【数2】

$$\frac{1}{\binom{n}{2}} \sum_{1 \leq i < j \leq n} \frac{\|f(x_i) - f(x_j)\|^2}{\|x_i - x_j\|^2},$$

式 1

$$E_{x, x' \sim P} \frac{\|f(x) - f(x')\|^2}{\|x - x'\|^2}$$

式 2

10

【0031】

[0035] 式2では、Eは、期待値を示し、Pは、 x sに関するデータ分布である。直感的に、この測度は、入力空間を拡張し、点を引き離す f の傾向を示す。より大きなメトリック表現性が、訓練後のより優れた性能と相関するはずである。この測度のバリエーションには、共通バッチのあらゆるペアを選択する代わりに (x, x') のペアをサンプリングし、2以外の指数を選択すること（すなわち、ある $p > 0$ に関して

【数3】

$$\frac{\|f(x_i) - f(x_j)\|^p}{\|x_i - x_j\|^p}$$

20

）、及びニューラルネットワークの「拡張性」（入力空間を拡張する傾向）をテストする他の方法が含まれる。

【0032】

[0036] メトリック表現性特徴に加えて、勾配変形は、最後の層に対するランダム勾配ベクトルをサンプリングし、各先行層で勾配ベクトルを取得するために各サンプル x_i に対して前方計算後に逆伝搬を行うことを含む。このような固定入力 x_i 及び固定の最後の層の勾配ベクトルを仮定して、パラメータ w に関する層1における勾配ベクトルが

30

【数4】

$$\left. \frac{\partial E}{\partial w^{(l)}} \right|_{x=x_i}$$

で示されるとする。その場合、パラメータ w の勾配変形は、式3で定義される。

【数5】

$$\frac{1}{n} \sum_{i=1}^n \left| \log \left\| \left. \frac{\partial E}{\partial w^{(l)}} \right|_{x=x_i} \right\| - \log \left\| \left. \frac{\partial E}{\partial w^{(1)}} \right|_{x=x_i} \right\| \right|$$

40

式 3

【0033】

[0037] つまり、勾配変形は、典型的なデータ点に対してどの程度勾配急増又は縮小が生じているかの測度である。これが大きくなるほど、訓練された性能が悪くなることが予期され、その理由は、（確率的勾配降下（SGD）による）訓練が難しいことが予期されるためである。この測度のバリエーションは、各 x_i に関して、新しい最後の層の勾配ベ

50

クトルをサンプリングすること、被加数を式 4 に置換すること、及び測定勾配急増 / 縮小の他の方法を含む。上記の測度は、固定アーキテクチャ及び初期化をランダム化した場合の残差ネットワーク性能の予測である。

【数 6】

$$\max\left(\left\|\frac{\partial E}{\partial w^{(L)}}\right\|_{x=x_i} / \left\|\frac{\partial E}{\partial w^{(1)}}\right\|_{x=x_1}, \left\|\frac{\partial E}{\partial w^{(1)}}\right\|_{x=x_1} / \left\|\frac{\partial E}{\partial w^{(L)}}\right\|_{x=x_1}\right)$$

式 4

10

【0034】

[0038] メトリック表現性の場合と同じセットアップを所与として、角度表現性が式 5 に定義される。

【数 7】

$$\frac{1}{\binom{n}{2}} \sum_{1 \leq i < j \leq n} \frac{f(x_i) \cdot f(x_j)}{\|f(x_i)\| \|f(x_j)\|}$$

式 5

20

【0035】

[0039] 角度表現性は、どの程度 f が、角度を引き離すという意味で、入力ベクトルを「非相関」させるかを測定する。従って、大きな角度表現性が、より優れた性能と相関することが予期される。幾つかのスキームでは、予測量は、実際に、 f の深さが無限に近づくときの漸近極限值からの

【数 8】

$$\frac{f(x_i) \cdot f(x_j)}{\|f(x_i)\| \|f(x_j)\|}$$

30

の偏差である。この漸近偏差の代理は、コーシー誤差 $C_1 - C(1+1)$ である (C_1 は、1 番目の層に至るまでのネットワークの角度表現性である)。

【0036】

[0040] 場合によっては、統計値は、自動的に学習され得る。本明細書に記載するデータベースの統計値は、モデルの最終的な性能と高度に相関するが、それらは、所与のモデルの一般化能力を予測するために抽出することができる最高予測統計値ではない可能性がある。最高予測統計値は、生の統計値 (すなわち、前のセクションにおける統計値を計算するために使用された埋め込み及び勾配測度) の複素非線形関数の可能性がある。

【0037】

[0041] それでもやはり、機械学習におけるその他の関数近似問題のように、十分な訓練データが存在すれば、これらの複素関数を学習することが可能となり得る。これは、サーバ 110 が生のデータから予測統計値も学習する別のバージョンのフレームワークの動機になるだろう。具体的には、幾つかの局面は、入力が最後の層におけるデータ表現 (すなわち、上記の $f(x)$)、最後の層及び最初の層の勾配などであるニューラルネットワークを使用し得る。

40

【0038】

[0042] 幾つかの局面の一般的バージョンでは、所定の所望の誤り率が達成されるまで、又は所定の総計算コストが達成されるまで、アルゴリズム 1 の手順が繰り返される。

アルゴリズム 1

1. 複数のディープニューラルネットワーク (DNN) アーキテクチャを提案する。D

50

NNアーキテクチャは、事前に定義されたアーキテクチャの空間からサンプリングされ、又は一般的なビルディングブロックを使用して構築され得る。

2. Glorot Normal初期化を使用して、アーキテクチャの重みを初期化する。代替実施形態において、独立正規分布及びラプラス分布が使用され得る。

3. アーキテクチャごとに、メトリック表現性及び勾配変形を計算する。

4. アーキテクチャ性能を予測するために、ベイジアンリカレントニューラルネットワーク（ベイジアンRNN）への入力として、各アーキテクチャのメトリック表現性及び勾配変形を含む。このタスクのために、ガウス過程を使用することもできる。原理上、予測に関する分布を出力するどのような確率的機械学習アルゴリズムも使用することができる。

5. expected improvementサンプリング、upper confidence boundサンプリング、又はトンプソンサンプリングなどの取得関数への入力として、アーキテクチャごとにベイジアンRNNを使用して計算された事後平均及び分散を使用する。

6. 閾値を超える取得関数の1つ又は複数の値を有する1つ又は複数のアーキテクチャをサンプリングする。

7. 閾値を超える取得関数の1つ又は複数の値を有する1つ又は複数のアーキテクチャを訓練し、提供セットに対するそれらの性能を評価する。観察された性能を用いて、動作4で使用されたモデルを更新する。動作1に戻る。

【0039】

[0043] アルゴリズム1は、図3に要約される。図3は、幾つかの実施形態による、誤り率を減少させる例示的方法のフローチャート300を示す。

【0040】

[0044] 動作310では、サーバ（例えば、サーバ110）が、複数のDNNアーキテクチャを提案する。DNNアーキテクチャは、事前に定義されたアーキテクチャの空間からサンプリングされ、又は一般的なビルディングブロックを使用して構築され得る。

【0041】

[0045] 動作320では、サーバが、Glorot Normal初期化を使用して、アーキテクチャの重みを初期化する。代替実施形態において、独立正規分布及びラプラス分布が使用され得る。

【0042】

[0046] 動作330では、サーバが、各アーキテクチャのメトリック表現性及び勾配変形を計算する。

【0043】

[0047] 動作340では、サーバが、アーキテクチャ性能を予測するために、ベイジアンRNNへの入力として、各アーキテクチャのメトリック表現性及び勾配変形を含む。このタスクのために、ガウス過程を使用することもできる。原理上、予測に関する分布を出力するどのような確率的機械学習アルゴリズムも使用することができる。

【0044】

[0048] 動作350では、サーバが、expected improvementサンプリング、upper confidence boundサンプリング、又はトンプソンサンプリングなどの取得関数への入力として、アーキテクチャごとにベイジアンRNNを使用して計算された事後平均及び分散を使用する。

【0045】

[0049] 動作360では、サーバが、閾値を超える取得関数の1つ又は複数の値を有する1つ又は複数のアーキテクチャをサンプリングする。

【0046】

[0050] 動作370では、サーバが、閾値を超える取得関数の1つ又は複数の値を有する1つ又は複数のアーキテクチャを訓練し、提供セットに対するそれらの性能を評価する。観察された性能に基づいて、動作340のベイジアンRNNモデルが更新される。動作370の後に、（動作360の）取得関数の値が十分であれば、方法300は終了する。そうでなければ、方法300は、動作310に戻る。

10

20

30

40

50

【 0 0 4 7 】

[0051] 具体的な一実施態様をここで説明する。回帰性基本ユニット（セル）が複数回繰り返される固定のディープニューラルネットワークアーキテクチャを所与として、幾つかの局面は、最高精度を有するセルを推論することに向けられる。まず、セルが、特定の数のブロックの有向非巡回グラフとして定義される。各ブロックは、2つの入力（ I_1 、 I_2 ）を取り込み、各入力に対して動作を行い（ O_1 、 O_2 ）、これら2つの動作からの出力の合計を返す。ブロックに関する1組の可能な入力は、セル内の全ての以前のブロックの出力及び以前の2つのセルの出力である。これらの入力に対する動作は、畳み込み、マックスプーリングなどの標準ニューラルネットワーク層である。ディープニューラルネットワークにおけるこのような基本ユニットの自動的且つ正確なサーチは、それらの性能が典型的には、より大きなデータセットへの適用において一般化するため重要である。B個のブロックを有する最良のセルを見つけるために、アルゴリズム2が使用される。

10

【 0 0 4 8 】

[0052] 場合によっては、アルゴリズム2におけるサーチは、1つのブロック（C）のセル数が小さくなり得るので、網羅的となり得る。例えば、8つの可能な動作の場合、サーチ空間は、 $8 * 8 * 2 * 2 = 256$ であり、これは、例えば256個のグラフィック処理ユニット（GPU）を並行して使用することによって、網羅的に探索され得る。

アルゴリズム2

1. 1つのブロック（C）における1組のセットを考える。

2. C内の各セルについて：

20

- ・Glorot Normal初期化を使用して、アーキテクチャの重みを初期化する。代替実施形態において、一様分布（i i d）の正規分布及びラプラス分布が使用され得る。

- ・メトリック表現性及び勾配変形を計算する。これは、複数の初期化にわたり行われ、平均及び分散が報告され得る。

- ・一定数のエポックの間、ニューラルネットワークを訓練する。

- ・テストセットに関する精度を計算する。

3. C内のセルのブロック数がBである場合、停止し、最高テスト精度を有するセルを返す。

4. セルアーキテクチャ及びランダム統計値を入力として取り込むベイジアンリカレントニューラルネットワーク（BRNN）を訓練し、テストセット精度を予測する。

30

5. C内の全てのセルを拡大し、新しいブロックを追加することによって取得されたセルを含むセットDを考える。

6. D内のセルに関してメトリック表現性及び勾配変形を計算する。この計算が（例えば、D内の多数のセルにより）高価すぎる場合には、（i）単にセル構造のみを取り込む追加のBRNNを訓練し、テストセット精度を予測することができ、（i i）それを使用して、予測されたテスト精度に基づいてセットDをフィルタリングすることができる。

7. 動作4のBRNNを使用して、次の反復で考えるセルのセット Dを決定する。この決定は、ベイジアン最適化フレームワーク内で利用と探索との間のトレードオフによって行われる。

8. Cならば、ステップ2に行く。

40

【 0 0 4 9 】

[0053] アルゴリズム2は、図4A～4Bに要約される。図4A～4Bは、幾つかの実施形態による、ブロックの中からセルを選択する例示的方法400のフローチャートを示す。

【 0 0 5 0 】

[0054] 図4Aに示されるように、動作410は、サブ動作412～418を含む。動作410では、サーバ（例えば、サーバ110）が、ブロックC内の各セルに関して、サブ動作412～418を実施する。

【 0 0 5 1 】

[0055] サブ動作412では、サーバが、Glorot Normal初期化を使用して、アーキテ

50

クチャの重みを初期化する。代替実施形態において、一様分布 (i i d) の正規分布及びラプラス分布が使用され得る。

【 0 0 5 2 】

[0056] サブ動作 4 1 4 では、サーバが、メトリック表現性及び勾配変形を計算する。これは、複数の初期化にわたり行われ、平均及び分散が報告され得る。

【 0 0 5 3 】

[0057] サブ動作 4 1 6 では、サーバが、一定数のエポックの間、ニューラルネットワークを訓練する。

【 0 0 5 4 】

[0058] サブ動作 4 1 8 では、サーバが、テストセットに関する精度を計算する。C 内の各セルに関して動作 4 1 0 が完了した後に、方法 4 0 0 は、動作 4 2 0 に進む。

10

【 0 0 5 5 】

[0059] 動作 4 2 0 では、サーバが、C のブロック数が所定の範囲内であるか否かを決定する。所定の範囲内であれば、方法 4 0 0 は、動作 4 3 0 に進む。所定の範囲内でなければ、方法 4 0 0 は、動作 4 4 0 に進む。

【 0 0 5 6 】

[0060] 動作 4 3 0 では、C のブロック数が所定の範囲内であれば、サーバは、最高テスト精度を有する C 内のセルを返す。動作 4 3 0 の後、方法 4 0 0 は終了する。

【 0 0 5 7 】

[0061] 動作 4 4 0 では、C のブロック数が所定の範囲内でなければ、サーバは、セルアーキテクチャ及びランダム統計値を入力として取り込む B R N N を訓練し、テストセット精度を予測する。動作 4 4 0 の後に、方法 4 0 0 は、図 4 B に示される動作 4 5 0 に進む。

20

【 0 0 5 8 】

[0062] 図 4 B に示すように、動作 4 5 0 では、サーバが、C 内の全てのセルを拡大し、新しいブロックを追加することによって取得されたセルを含むセット D を生成する。

【 0 0 5 9 】

[0063] 動作 4 6 0 では、サーバが、D 内のセルに関してメトリック表現性及び勾配変形を計算する。この計算が (例えば、D 内の多数のセルにより) 高価すぎる場合には、サーバは、(i) 単にセル構造のみを取り込む追加の B R N N を訓練し、テストセット精度を予測し、(i i) この追加の B R N N を使用して、予測されたテスト精度に基づいてセット D をフィルタリングし得る。

30

【 0 0 6 0 】

[0064] 動作 4 7 0 では、サーバが、動作 4 4 0 で訓練された B R N N を使用して、次の反復で考えるセルのセット D を決定する。この決定は、ベイジアン最適化フレームワーク内で利用と探索との間のトレードオフによって行われる。

【 0 0 6 1 】

[0065] 動作 4 8 0 では、サーバが、 を C として保存する。動作 4 8 0 の後に、方法 4 0 0 は、図 4 A の動作 4 2 0 に戻る。

【実施例】

40

【 0 0 6 2 】

番号付き実施例

[0066] 番号付き実施例 1、2、3 などとして、特定の実施形態をここで説明する。これらの番号付き実施例は、単なる例として提供されるものであり、主題技術を限定するものではない。

【 0 0 6 3 】

[0067] 実施例 1 は、処理ハードウェア、並びに機械学習問題と関連付けられた機械学習問題空間、及び機械学習問題を解くための複数の訓練を受けていない候補ニューラルネットワークにアクセスすることと、訓練を受けていない候補ニューラルネットワークごとに、機械学習問題に対する候補ニューラルネットワークの表現性を捕捉する少なくとも 1

50

つの表現性測度を計算することと、訓練を受けていない候補ニューラルネットワークごとに、機械学習問題に対する候補ニューラルネットワークの訓練可能性を捕捉する少なくとも1つの訓練可能性測度を計算することと、少なくとも1つの表現性測度及び少なくとも1つの訓練可能性測度に基づいて、機械学習問題を解くための少なくとも1つの候補ニューラルネットワークを選択することと、選択された少なくとも1つの候補ニューラルネットワークを表す出力を提供することと、を含む動作を処理ハードウェアに行わせる命令を保存するメモリを含むシステムである。

【0064】

[0068] 実施例2では、実施例1の特定事項は、少なくとも1つの表現性測度が、訓練を受けていない候補ニューラルネットワークによる、機械学習問題空間からのサンプルの分離の測度を表すことを含む。

10

【0065】

[0069] 実施例3では、実施例2の特定事項は、分離の測度が大きさであることを含む。

【0066】

[0070] 実施例4では、実施例2～3の特定事項は、分離の測度が角度であることを含む。

【0067】

[0071] 実施例5では、実施例1～4の特定事項は、少なくとも1つの訓練可能性測度が、訓練の第1段階の間、候補ニューラルネットワークにおける重みの確率的勾配降下を表すことを含む。

20

【0068】

[0072] 実施例6では、実施例1～5の特定事項は、機械学習問題を解くための少なくとも1つの候補ニューラルネットワークを選択することが、閾値を超える少なくとも1つの表現性測度、及び範囲内の少なくとも1つの訓練可能性測度を有する少なくとも1つの候補ニューラルネットワークを選択することを含み、この範囲が、範囲最小値及び範囲最大値によって定義されることを含む。

【0069】

[0073] 実施例7では、実施例1～6の特定事項は、動作が、機械学習問題を解くために少なくとも1つの候補ニューラルネットワークを訓練することをさらに含むことを含む。

【0070】

30

[0074] 実施例8では、実施例7の特定事項は、動作が、機械学習問題を解くために、訓練された少なくとも1つの候補ニューラルネットワークを機械学習問題空間で実行することと、訓練された少なくとも1つの候補ニューラルネットワークによって生成された機械学習問題に対する解を提供することと、をさらに含むことを含む。

【0071】

[0075] 実施例9は、機械学習問題と関連付けられた機械学習問題空間、及び機械学習問題を解くための複数の訓練を受けていない候補ニューラルネットワークにアクセスすることと、訓練を受けていない候補ニューラルネットワークごとに、機械学習問題に対する候補ニューラルネットワークの表現性を捕捉する少なくとも1つの表現性測度を計算することと、訓練を受けていない候補ニューラルネットワークごとに、機械学習問題に対する候補ニューラルネットワークの訓練可能性を捕捉する少なくとも1つの訓練可能性測度を計算することと、少なくとも1つの表現性測度及び少なくとも1つの訓練可能性測度に基づいて、機械学習問題を解くための少なくとも1つの候補ニューラルネットワークを選択することと、選択された少なくとも1つの候補ニューラルネットワークを表す出力を提供することと、を含む動作を1つ又は複数の機械に行わせる命令を保存する、非一時的機械可読媒体である。

40

【0072】

[0076] 実施例10では、実施例9の特定事項は、少なくとも1つの表現性測度が、訓練を受けていない候補ニューラルネットワークによる、機械学習問題空間からのサンプルの分離の測度を表すことを含む。

50

【 0 0 7 3 】

[0077] 実施例 1 1 では、実施例 1 0 の特定事項は、分離の測度が大きさであることを含む。

【 0 0 7 4 】

[0078] 実施例 1 2 では、実施例 1 0 ~ 1 1 の特定事項は、分離の測度が角度であることを含む。

【 0 0 7 5 】

[0079] 実施例 1 3 では、実施例 9 ~ 1 2 の特定事項は、少なくとも 1 つの訓練可能性測度が、訓練の第 1 段階の間、候補ニューラルネットワークにおける重みの確率的勾配降下を表すことを含む。

10

【 0 0 7 6 】

[0080] 実施例 1 4 では、実施例 9 ~ 1 3 の特定事項は、機械学習問題を解くための少なくとも 1 つの候補ニューラルネットワークを選択することが、閾値を超える少なくとも 1 つの表現性測度、及び範囲内の少なくとも 1 つの訓練可能性測度を有する少なくとも 1 つの候補ニューラルネットワークを選択することを含み、この範囲が、範囲最小値及び範囲最大値によって定義されることを含む。

【 0 0 7 7 】

[0081] 実施例 1 5 は、機械学習問題と関連付けられた機械学習問題空間、及び機械学習問題を解くための複数の訓練を受けていない候補ニューラルネットワークにアクセスすることと、訓練を受けていない候補ニューラルネットワークごとに、機械学習問題に対する候補ニューラルネットワークの表現性を捕捉する少なくとも 1 つの表現性測度を計算することと、訓練を受けていない候補ニューラルネットワークごとに、機械学習問題に対する候補ニューラルネットワークの訓練可能性を捕捉する少なくとも 1 つの訓練可能性測度を計算することと、少なくとも 1 つの表現性測度及び少なくとも 1 つの訓練可能性測度に基づいて、機械学習問題を解くための少なくとも 1 つの候補ニューラルネットワークを選択することと、選択された少なくとも 1 つの候補ニューラルネットワークを表す出力を提供することと、を含む方法である。

20

【 0 0 7 8 】

[0082] 実施例 1 6 では、実施例 1 5 の特定事項は、少なくとも 1 つの表現性測度が、訓練を受けていない候補ニューラルネットワークによる、機械学習問題空間からのサンプルの分離の測度を表すことを含む。

30

【 0 0 7 9 】

[0083] 実施例 1 7 では、実施例 1 6 の特定事項は、分離の測度が大きさであることを含む。

【 0 0 8 0 】

[0084] 実施例 1 8 では、実施例 1 6 ~ 1 7 の特定事項は、分離の測度が角度であることを含む。

【 0 0 8 1 】

[0085] 実施例 1 9 では、実施例 1 5 ~ 1 8 の特定事項は、少なくとも 1 つの訓練可能性測度が、訓練の第 1 段階の間、候補ニューラルネットワークにおける重みの確率的勾配降下を表すことを含む。

40

【 0 0 8 2 】

[0086] 実施例 2 0 では、実施例 1 5 ~ 1 9 の特定事項は、機械学習問題を解くための少なくとも 1 つの候補ニューラルネットワークを選択することが、閾値を超える少なくとも 1 つの表現性測度、及び範囲内の少なくとも 1 つの訓練可能性測度を有する少なくとも 1 つの候補ニューラルネットワークを選択することを含み、この範囲が、範囲最小値及び範囲最大値によって定義されることを含む。

【 0 0 8 3 】

[0087] 実施例 2 1 は、処理回路網によって実行されると、実施例 1 ~ 2 0 の何れかを実施するための動作を処理回路網に行わせる命令を含む少なくとも 1 つの機械可読媒体で

50

ある。

【0084】

[0088] 実施例22は、実施例1～20の何れかを実施するための手段を含む装置である。

【0085】

[0089] 実施例23は、実施例1～20の何れかを実施するためのシステムである。

【0086】

[0090] 実施例24は、実施例1～20の何れかを実施するための方法である。

【0087】

コンポーネント及び論理

[0091] 特定の実施形態は、論理又は幾つかのコンポーネント若しくは機構を含むと本明細書に記載される。コンポーネントは、ソフトウェアコンポーネント（例えば、機械可読媒体で具現化されるコード）又はハードウェアコンポーネントを構成し得る。「ハードウェアコンポーネント」は、特定の動作を行うことが可能な有形ユニットであり、特定の物理的様式で構成又は配置され得る。様々な例示の実施形態では、1つ若しくは複数のコンピュータシステム（例えば、スタンドアロンコンピュータシステム、クライアントコンピュータシステム、若しくはサーバコンピュータシステム）、又はコンピュータシステムの1つ若しくは複数のハードウェアコンポーネント（例えば、プロセッサ若しくはプロセッサ群）が、ソフトウェア（例えば、アプリケーション又はアプリケーション部分）によって、本明細書に記載するような特定の動作を行うために動作するハードウェアコンポーネントとして構成され得る。

【0088】

[0092] 幾つかの実施形態では、ハードウェアコンポーネントは、機械的に、電子的に、又はそれらの任意の適宜の組み合わせで実施され得る。例えば、ハードウェアコンポーネントは、特定の動作を行うように恒久的に構成された専用回路網又は論理を含み得る。例えば、ハードウェアコンポーネントは、フィールドプログラマブルゲートアレイ（FPGA）又は特定用途向け集積回路（ASIC）などの専用プロセッサでもよい。ハードウェアコンポーネントは、特定の動作を行うようにソフトウェアによって一時的に構成されたプログラマブル論理又は回路網も含み得る。例えば、ハードウェアコンポーネントは、汎用プロセッサ又は他のプログラマブルプロセッサによって実行されるソフトウェアを含み得る。一旦このようなソフトウェアによって構成されると、ハードウェアコンポーネントは、構成された機能を行うように独自に調整された、もはや汎用プロセッサではない専用機械（又は機械の専用コンポーネント）となる。ハードウェアコンポーネントを機械的に、専用且つ恒久的に構成された回路網で、又は（例えば、ソフトウェアによって構成された）一時的に構成された回路網で実施するかの決定は、コスト及び時間を考慮することによって動かされ得ることが理解されるだろう。

【0089】

[0093] 従って、「ハードウェアコンポーネント」というフレーズは、有形レコードを包含する（ある特定の様式で動作する、又は本明細書に記載する特定の動作を行うように、物理的に構築された、恒久的に構成された（例えば、配線接続された）、又は一時的に構成された（例えば、プログラムされた）レコードである）と理解されるものとする。本明細書では、「ハードウェア実施コンポーネント」は、ハードウェアコンポーネントを指す。ハードウェアコンポーネントが一時的に構成される（例えば、プログラムされる）実施形態を考慮すると、各ハードウェアコンポーネントは、時間のあるインスタンスにおいて、構成されていない、又はインスタンス化されていない可能性がある。例えば、ハードウェアコンポーネントが、専用プロセッサとなるようにソフトウェアによって構成された汎用プロセッサを含む場合、この汎用プロセッサは、異なる時点で、それぞれ異なる専用プロセッサ（例えば、異なるハードウェアコンポーネントを含む）として構成され得る。従って、ソフトウェアは、例えば、時間のあるインスタンスにおいて、ある特定のハードウェアコンポーネントを構成し、異なる時間のインスタンスにおいて、異なるハードウェア

10

20

30

40

50

アコンポーネントを構成するように、特定の1つ又は複数のプロセッサを構成する。

【0090】

[0094] ハードウェアコンポーネントは、他のハードウェアコンポーネントに情報を提供し、及び他のハードウェアコンポーネントから情報を受信することができる。従って、記載したハードウェアコンポーネントは、通信可能に結合されると見なすことができる。複数のハードウェアコンポーネントが同時に存在する場合、通信は、ハードウェアコンポーネントの2つ以上の間の（例えば、適切な回路及びバスを利用した）信号伝送により達成され得る。複数のハードウェアコンポーネントが、異なる時点で構成又はインスタンス化される実施形態では、このようなハードウェアコンポーネント間の通信は、例えば、複数のハードウェアコンポーネントがアクセスしたメモリ構造における情報の保存及び取り出しにより達成され得る。例えば、あるハードウェアコンポーネントは、ある動作を行い、それが通信可能に結合したメモリデバイスにその動作の出力を保存し得る。その場合、別のハードウェアコンポーネントが、後に、保存された出力を取り出し、及び処理するために、そのメモリデバイスにアクセスし得る。ハードウェアコンポーネントは、入力デバイス又は出力デバイスとの通信も開始することができ、リソース（例えば、情報の集まり）を操作することができる。

10

【0091】

[0095] 本明細書に記載する例示的方法の様々な動作は、少なくとも部分的に、関連する動作を行うように（例えば、ソフトウェアによって）一時的に構成された、又は恒久的に構成された1つ又は複数のプロセッサによって行われ得る。一時的に構成されたものであれ、恒久的に構成されたものであれ、このようなプロセッサは、本明細書に記載する1つ又は複数の動作又は機能を行うように動作するプロセッサ実施コンポーネントを構成し得る。本明細書では、「プロセッサ実施コンポーネント」は、1つ又は複数のプロセッサを使用して実施されるハードウェアコンポーネントを指す。

20

【0092】

[0096] 同様に、本明細書に記載する方法は、特定の1つ又は複数のプロセッサがハードウェアの一例である状態で、少なくとも部分的にプロセッサ実施され得る。例えば、方法の動作の少なくとも幾つかは、1つ又は複数のプロセッサ又はプロセッサ実施コンポーネントによって行われ得る。また、1つ又は複数のプロセッサは、「クラウドコンピューティング」環境における関連動作の遂行を支援するために、又は「サービスとしてのソフトウェア」（SaaS）としても動作し得る。例えば、動作の少なくとも幾つかは、これらの動作が、ネットワーク（例えば、インターネット）を介して、及び1つ又は複数の適切なインタフェース（例えば、API）を介してアクセス可能な状態で、コンピュータ群（プロセッサを含む機械の例として）によって行われ得る。

30

【0093】

[0097] 動作の幾つかの遂行は、単一の機械内に常駐するだけでなく、幾つかの機械にわたって配備されたプロセッサ間で分散され得る。幾つかの例示的实施形態では、プロセッサ又はプロセッサ実施コンポーネントは、単一の地理的な場所（例えば、自宅環境、オフィス環境、又はサーバファーム内）に位置し得る。他の例示的实施形態では、プロセッサ又はプロセッサ実施コンポーネントは、幾つかの地理的な場所にわたって分散され得る。

40

【0094】

例示的機械及びソフトウェアアーキテクチャ

[0098] 図1～4に関連して記載されるコンポーネント、方法、アプリケーションなどは、幾つかの実施形態において、機械及び関連付けられたソフトウェアアーキテクチャの文脈で実施される。以下のセクションは、開示の実施形態での使用に適した代表的な1つ又は複数のソフトウェアアーキテクチャ及び1つ又は複数の機械（例えば、ハードウェア）アーキテクチャを説明する。

【0095】

[0099] ソフトウェアアーキテクチャは、特定の目的に合わせたデバイス及び機械を製作するために、ハードウェアアーキテクチャと併せて使用される。例えば、ある特定のソ

50

ソフトウェアアーキテクチャと結合された、ある特定のハードウェアアーキテクチャは、携帯電話、タブレットデバイスなどのモバイル機器を製作する。少し異なるハードウェア及びソフトウェアアーキテクチャは、「インターネットオブシングス」で使用するスマートデバイスを生み出すことができ、さらに別の組み合わせが、クラウドコンピューティングアーキテクチャ内で使用するサーバコンピュータを生じさせる。本明細書に含まれる開示内容とは異なる文脈で開示される特定事項を実施する方法を当業者は容易に理解することができるため、このようなソフトウェアアーキテクチャ及びハードウェアアーキテクチャの全ての組み合わせがここに提示されるわけではない。

【0096】

[00100] 図5は、機械可読媒体（例えば、機械可読ストレージ媒体）から命令を読み出し、本明細書で述べた方法論の1つ又は複数を行うことが可能な、幾つかの例示的实施形態による機械500のコンポーネントを示すブロック図である。具体的には、図5は、本明細書で述べた方法論の1つ又は複数を実行させる命令516（例えば、ソフトウェア、プログラム、アプリケーション、アプレット、アプリ、又は他の実行可能コード）が内部で実行され得るコンピュータシステムの例示的形式の機械500の図表示を示す。命令516は、一般的なプログラムされていない機械を、記載した様式で記載及び図示した機能を実行するようにプログラムされた特定の機械に変える。代替実施形態では、機械500は、スタンドアロンデバイスとして動作し、又は他の機械と結合（例えば、ネットワーク化）され得る。ネットワーク化された配備では、機械500は、サーバ-クライアントネットワーク環境におけるサーバ機若しくはクライアント機として、又はピアツーピア（若しくは分散型）ネットワーク環境におけるピア機として動作し得る。機械500には、サーバコンピュータ、クライアントコンピュータ、PC、タブレットコンピュータ、ラップトップコンピュータ、ネットブック、携帯情報端末（PDA）、娯楽メディアシステム、セルラー電話、スマートフォン、モバイル機器、ウェアラブルデバイス（例えば、スマートウォッチ）、スマートホームデバイス（例えば、スマート家電）、他のスマートデバイス、ウェブアプライアンス、ネットワークルータ、ネットワークスイッチ、ネットワークブリッジ、又は機械500が取るべきアクションを指定する命令516を順次又は別様に実行可能なあらゆる機械が含まれ得るが、これらに限定されない。さらに、単一の機械500のみが示されているが、「機械」という用語は、本明細書で述べる方法論の1つ又は複数を行うための命令516を個々に又は一緒に実行する機械500の集まりを含むとも解釈されるものとする。

【0097】

[00101] 機械500は、バス502などを介して互いに通信するように構成され得る、プロセッサ510、メモリ/ストレージ530、及び入出力コンポーネント550を含み得る。ある例示的实施形態では、プロセッサ510（例えば、中央処理装置（CPU）、縮小命令セットコンピューティング（RISC）プロセッサ、複数命令セットコンピューティング（CISC）プロセッサ、グラフィック処理ユニット（GPU）、デジタル信号プロセッサ（DSP）、ASIC、無線周波数集積回路（RFIC）、別のプロセッサ、又はこれらの任意の適宜の組み合わせ）は、例えば、命令516を実行し得るプロセッサ512及びプロセッサ514を含み得る。「プロセッサ」という用語は、命令を同時に実行し得る2つ以上の独立したプロセッサ（「コア」と呼ばれることもある）を含み得るマルチコアプロセッサを含むよう意図されている。図5は、複数のプロセッサ510を示しているが、機械500は、単一コアを備えた単一プロセッサ、複数のコアを備えた単一プロセッサ（例えば、マルチコアプロセッサ）、単一コアを備えた複数のプロセッサ、複数のコアを備えた複数のプロセッサ、又はこれらの任意の組み合わせを含み得る。

【0098】

[00102] メモリ/ストレージ530は、メインメモリ又は他のメモリストレージなどのメモリ532、及びストレージユニット536（両方とも、バス502などを介してプロセッサ510にアクセス可能である）を含み得る。ストレージユニット536及びメモリ532は、本明細書に記載する方法論又は機能の1つ又は複数を実現化する命令516

10

20

30

40

50

を保存する。命令 5 1 6 はまた、機械 5 0 0 によるこれらの実行中に、メモリ 5 3 2 内、ストレージユニット 5 3 6 内、プロセッサ 5 1 0 の少なくとも 1 つの内部（例えば、プロセッサのキャッシュメモリ内）、又はこれらの任意の適宜の組み合わせに完全又は部分的に存在し得る。従って、メモリ 5 3 2、ストレージユニット 5 3 6、及びプロセッサ 5 1 0 のメモリは、機械可読媒体の例である。

【 0 0 9 9 】

[00103] 本明細書では、「機械可読媒体」は、命令（例えば、命令 5 1 6）及びデータを一時的又は恒久的に保存可能なデバイスを意味し、ランダムアクセスメモリ（RAM）、読み出し専用メモリ（ROM）、バッファメモリ、フラッシュメモリ、光媒体、磁気媒体、キャッシュメモリ、他のタイプのストレージ（例えば、消去可能プログラマブル読み出し専用メモリ（EEPROM））、及び/又はこれらの任意の適宜の組み合わせを含み得るが、これらに限定されない。「機械可読媒体」という用語は、命令 5 1 6 を保存することができる単一の媒体又は複数の媒体（例えば、集中型若しくは分散型データベース、又は関連付けられたキャッシュ及びサーバ）を含むと解釈されるものとする。「機械可読媒体」という用語はまた、機械（例えば、機械 5 0 0）によって実行される命令（例えば、命令 5 1 6）を保存する（命令が機械の 1 つ又は複数のプロセッサ（例えば、プロセッサ 5 1 0）によって実行されると、本明細書に記載する方法論の 1 つ又は複数を行わせるように）ことが可能なあらゆる媒体又は複数の媒体のあらゆる組み合わせを含むと解釈されるものとする。従って、「機械可読媒体」は、単一のストレージ装置又はデバイス、及び複数のストレージ装置又はデバイスを含む「クラウドベースの」ストレージシステム又はストレージネットワークを指す。「機械可読媒体」という用語は、正確な意味において、信号を除外する。

【 0 1 0 0 】

[00104] 入出力コンポーネント 5 5 0 は、入力を受け取り、出力を提供し、出力を生成し、情報を送信し、情報をやり取りし、測定値を捕捉するなどのための多様なコンポーネントを含み得る。ある特定の機械に含まれる具体的な入出力コンポーネント 5 5 0 は、機械のタイプに左右される。例えば、携帯電話などのポータブル機は、恐らく、タッチ入力デバイス又は他のそのような入力機構を含むが、ヘッドレスサーバ機は、恐らく、そのようなタッチ入力デバイスを含まない。入出力コンポーネント 5 5 0 が、図 5 に示されない多くの他のコンポーネントを含み得ることが理解されるだろう。入出力コンポーネント 5 5 0 は、以下の説明を簡単にするためだけに機能性に基づいてグループ化され、このグループ化は、決して制限するものではない。様々な例示的实施形態において、入出力コンポーネント 5 5 0 は、出力コンポーネント 5 5 2 及び入力コンポーネント 5 5 4 を含み得る。出力コンポーネント 5 5 2 は、視覚的コンポーネント（例えば、プラズマディスプレイパネル（PDP）、発光ダイオード（LED）ディスプレイ、液晶ディスプレイ（LCD）、プロジェクタ、又は陰極線管（CRT））、音響コンポーネント（例えば、スピーカ）、ハプティックコンポーネント（例えば、振動モータ、抵抗機構）、他の信号生成器などを含み得る。入力コンポーネント 5 5 4 は、英数字入力コンポーネント（例えば、キーボード、英数字入力を受けるように構成されたタッチスクリーン、光学式キーボード、又は他の英数字入力コンポーネント）、ポイントベースの入力コンポーネント（例えば、マウス、タッチパッド、トラックボール、ジョイスティック、運動センサ、又は別のポインティング機器）、触覚入力コンポーネント（例えば、物理的ボタン、タッチ若しくはタッチジェスチャの場所及び/又は力を提供するタッチスクリーン、又は他の触覚入力コンポーネント）、音声入力コンポーネント（例えば、マイクロフォン）などを含み得る。

【 0 1 0 1 】

[00105] さらに例示的实施形態では、入出力コンポーネント 5 5 0 は、多様な他のコンポーネントの中でも、バイオメトリックコンポーネント 5 5 6、運動コンポーネント 5 5 8、環境コンポーネント 5 6 0、又は位置コンポーネント 5 6 2 を含み得る。例えば、バイオメトリックコンポーネント 5 5 6 は、表現（例えば、手振り、表情、声による表現、身振り、又は視線追跡）を検出し、生体信号（例えば、血圧、心拍数、体温、発汗、

10

20

30

40

50

又は脳波)を測定し、運動関連のメトリック(例えば、移動距離、移動の速度、又は運動に費やした時間)を測定し、人を識別する(例えば、声の識別、網膜識別、顔識別、指紋識別、又は脳波図ベースの識別)などのためのコンポーネントを含み得る。運動コンポーネント558は、加速度センサコンポーネント(例えば、加速度計)、重力センサコンポーネント、回転センサコンポーネント(例えば、ジャイロスコープ)などを含み得る。環境コンポーネント560は、例えば、照明センサコンポーネント(例えば、光度計)、温度センサコンポーネント(例えば、周囲温度を検出する1つ若しくは複数の温度計)、湿度センサコンポーネント、圧力センサコンポーネント(例えば、圧力計)、音響センサコンポーネント(例えば、背景雑音を検出する1つ若しくは複数のマイクロフォン)、近接センサコンポーネント(例えば、近くの物体を検出する赤外線センサ)、ガスセンサ(例えば、安全のために有害ガスの濃度を検出するため、若しくは大気中の汚染物質を測定するためのガス検出センサ)、又は周囲の物理的環境に対応する表示、測定値、又は信号を提供し得る他のコンポーネントを含み得る。位置コンポーネント562は、場所センサコンポーネント(例えば、全地球測位システム(GPS)受信機コンポーネント)、高度センサコンポーネント(例えば、それから高度が導出され得る気圧を検出する高度計又は圧力計)、方位センサコンポーネント(例えば、磁力計)などを含み得る。

10

【0102】

[00106] 通信は、多様な技術を用いて実施され得る。入出力コンポーネント550は、結合582及び結合572によりネットワーク580又はデバイス570に機械500をそれぞれ結合させるように動作可能な通信コンポーネント564を含み得る。例えば、通信コンポーネント564は、ネットワーク580とインタフェースを取るネットワークインタフェースコンポーネント又は他の適宜のデバイスを含み得る。さらなる例では、通信コンポーネント564は、有線通信コンポーネント、無線通信コンポーネント、セルラー通信コンポーネント、近距離無線通信(NFC)コンポーネント、Bluetooth(登録商標)コンポーネント(例えば、Bluetooth(登録商標)Low Energy)、Wi-Fi(登録商標)コンポーネント、及び他のモダリティによる通信を提供するための他の通信コンポーネントを含み得る。デバイス570は、別の機械又は多様な周辺デバイスの何れか(例えば、USBを介して結合された周辺デバイス)でもよい。

20

【0103】

[00107] また、通信コンポーネント564は、識別子を検出することができ、又は識別子を検出するように動作可能なコンポーネントを含んでもよい。例えば、通信コンポーネント564は、無線認識(RFID)タグ読取り機コンポーネント、NFCスマートタグ検出コンポーネント、光学式読取り機コンポーネント、又は音響検出コンポーネント(例えば、タグを付けた音声信号を識別するためのマイクロフォン)を含み得る。さらに、通信コンポーネント564を用いて様々な情報を導出することができる(インターネットプロトコル(IP)ジオロケーションによる場所、Wi-Fi(登録商標)信号三角測量による場所、ある特定の場所を示し得るNFCビーコン信号を検出することによる場所など)。

30

【0104】

[00108] 様々な例示的实施形態では、ネットワーク580の1つ又は複数の部分は、アドホックネットワーク、イントラネット、エクストラネット、仮想プライベートネットワーク(VPN)、ローカルエリアネットワーク(LAN)、無線LAN(WLAN)、WAN、無線WAN(WWAN)、メトロポリタンエリアネットワーク(MAN)、インターネット、インターネットの一部、公衆交換電話網(PSTN)の一部、単純旧式電話サービス(POTS)ネットワーク、セルラー電話ネットワーク、無線ネットワーク、Wi-Fi(登録商標)ネットワーク、別のタイプのネットワーク、又は2つ以上の上記ネットワークの組み合わせでもよい。例えば、ネットワーク580又はネットワーク580の一部は、無線又はセルラーネットワークを含んでもよく、結合582は、符号分割多重アクセス(CDMA)接続、グローバルシステムフォーモバイルコミュニケーションズ(GSM)接続、又は別のタイプのセルラー若しくは無線結合でもよい。この例では、結合582は、シングルキャリア無線伝送技術(1xRTT)、EVDO(Evolution-Data Optimi

40

50

zed) 技術、汎用パケット無線サービス (GPRS) 技術、GSM 進化型高速データレート (EDGE) 技術、5G を含む第3世代パートナーシッププロジェクト (3GPP)、第4世代無線 (4G) ネットワーク、ユニバーサル移動体通信システム (UMTS)、高速パケットアクセス (HSPA)、ワイマックス (WiMAX: Worldwide Interoperability for Microwave Access)、ロングタームエボリューション (LTE) 規格、様々な規格設定団体によって定義された他の規格、他の長距離プロトコル、又は他のデータ転送技術などの様々なタイプのデータ転送技術の何れかを実施し得る。

【0105】

[00109] 命令516は、ネットワークインタフェースデバイス (例えば、通信コンポーネント564に含まれるネットワークインタフェースコンポーネント) により、伝送媒体を使用し、幾つかの周知の転送プロトコル (例えばHTTP) の何れか1つを利用してネットワーク580上で送信又は受信が行われ得る。同様に、命令516は、デバイス570への結合572 (例えば、ピアツーピア結合) により、伝送媒体を使用して送信又は受信が行われ得る。「伝送媒体」という用語は、機械500による実行のために命令516を保存し、符号化し、又は運ぶことが可能なあらゆる無形媒体を含むと解釈されるものとし、このようなソフトウェアの通信を容易にするデジタル若しくはアナログ通信信号、又は他の無形媒体を含む。

10

20

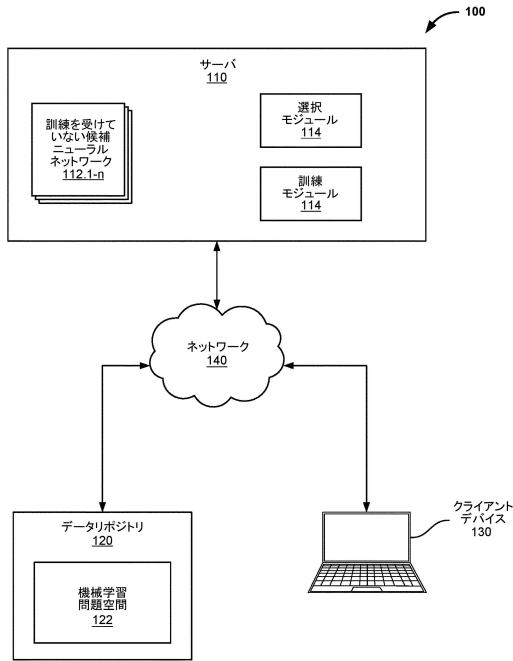
30

40

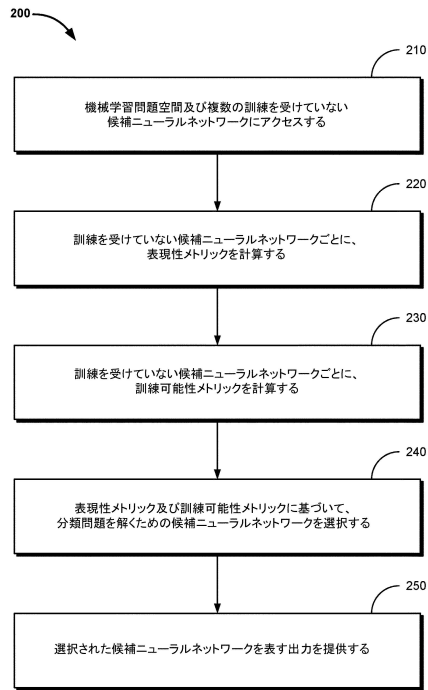
50

【図面】

【図 1】



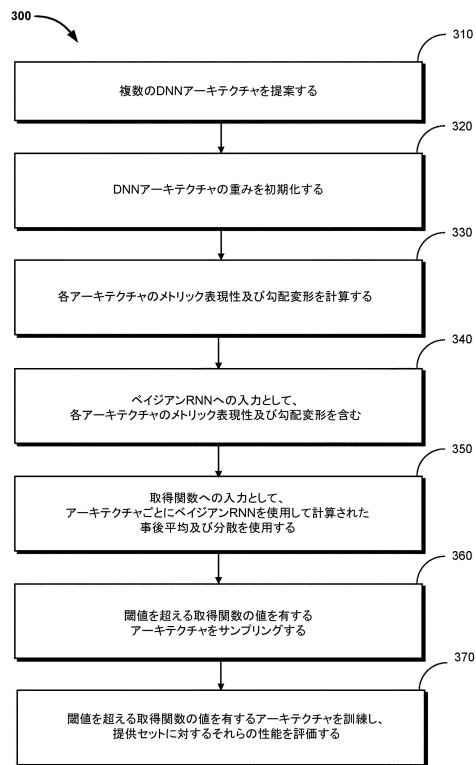
【図 2】



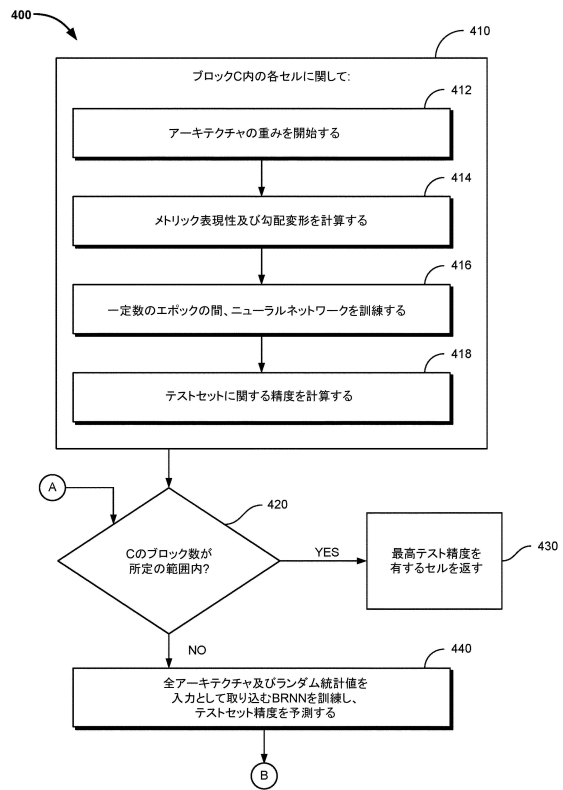
10

20

【図 3】



【図 4 A】

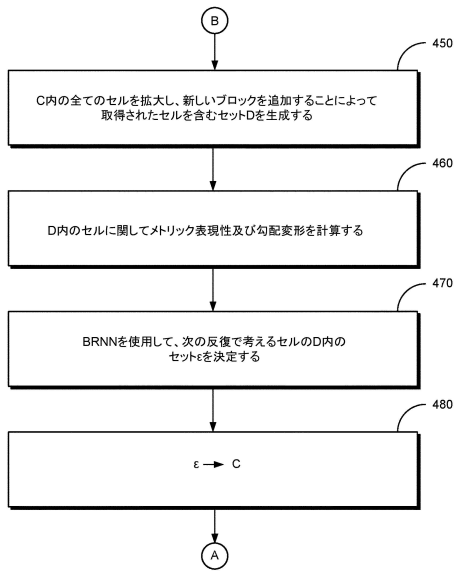


30

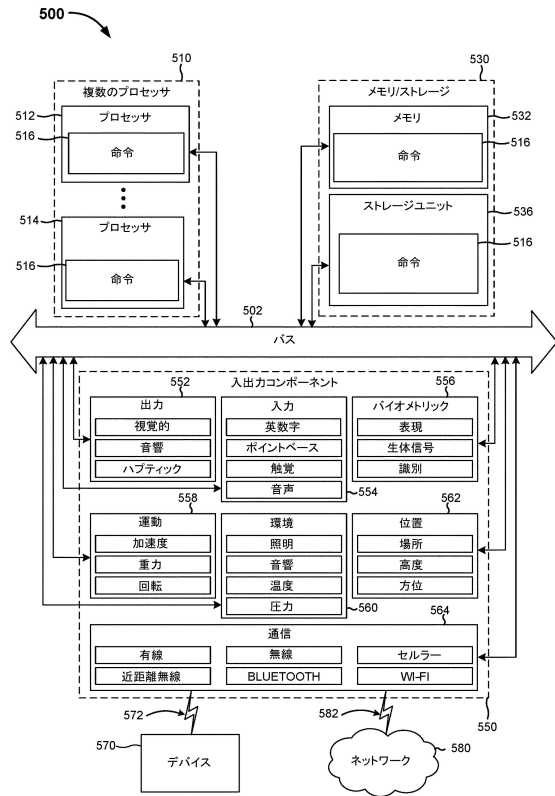
40

50

【 図 4 B 】



【 図 5 】



フロントページの続き

- 弁理士 阿部 豊隆
 (74)代理人 100134027
 弁理士 松田 達也
 (72)発明者 アミザデ, サイード
 アメリカ合衆国, ワシントン州 98052-6399, レッドモンド, ワン マイクロソフト ウ
 ェイ, マイクロソフト テクノロジー ライセンシング, エルエルシー
 (72)発明者 ヤン, ゲ
 アメリカ合衆国, ワシントン州 98052-6399, レッドモンド, ワン マイクロソフト ウ
 ェイ, マイクロソフト テクノロジー ライセンシング, エルエルシー
 (72)発明者 フシ, ニコロ
 アメリカ合衆国, ワシントン州 98052-6399, レッドモンド, ワン マイクロソフト ウ
 ェイ, マイクロソフト テクノロジー ライセンシング, エルエルシー
 (72)発明者 カザーレ, フランチェスコ パオロ
 アメリカ合衆国, ワシントン州 98052-6399, レッドモンド, ワン マイクロソフト ウ
 ェイ, マイクロソフト テクノロジー ライセンシング, エルエルシー
 審査官 松平 英
 (56)参考文献 特表2017-520068(JP, A)
 特開平05-314090(JP, A)
 特表2015-533437(JP, A)
 米国特許出願公開第2015/0339572(US, A1)
 国際公開第2017/058489(WO, A1)
 Andrew Brock et al, SMASH: One-Shot Model Architecture Search through HyperNetworks
 , [online], 2017年08月17日, [検索日 2023年3月28日]、インターネット<<http://arxiv.org/abs/1708.05344>>
 Philippe Hamel et al, TRANSFER LEARNING IN MIR: SMARING LEARNED LATENT REPRES
 ENTATIONS FOR MUSIC AUDIO CLASSIFICATION AND SIMILARITY, [online], 2013年,
 [検索日 2023年3月28日]、インターネット<<https://archives.ismir.net/ismir2013/paper/000076.pdf>>
 Thomas Elsken et al, SIMPLE AND EFFICIENT ARCHITECTURE SEARCH FOR CONVOLUTI
 ONAL NEURAL NETWORKS, [online], 2017年11月13日, [検索日 2023年3月28
 日]、インターネット<<https://arxiv.org/abs/1711.04528>>
 (58)調査した分野 (Int.Cl., DB名)
 G06N 3/00 - 3/12
 7/08 - 99/00
 G06T 7/00 - 7/90
 G06V 10/00 - 20/90
 30/418
 40/16
 40/20