

(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号
特許第7702016号
(P7702016)

(45)発行日 令和7年7月2日(2025.7.2)

(24)登録日 令和7年6月24日(2025.6.24)

(51)国際特許分類 F I
 G 0 6 N 20/00 (2019.01) G 0 6 N 20/00
 G 0 6 F 11/36 (2025.01) G 0 6 F 11/3668 1 8 8

請求項の数 20 外国語出願 (全16頁)

(21)出願番号	特願2024-66930(P2024-66930)	(73)特許権者	521042770 ウーブン・バイ・トヨタ株式会社 東京都中央区日本橋室町三丁目2番1号
(22)出願日	令和6年4月17日(2024.4.17)	(74)代理人	100099759 弁理士 青木 篤
(65)公開番号	特開2024-180284(P2024-180284 A)	(74)代理人	100123582 弁理士 三橋 真二
(43)公開日	令和6年12月26日(2024.12.26)	(74)代理人	100092624 弁理士 鶴田 準一
審査請求日	令和6年4月17日(2024.4.17)	(74)代理人	100147555 弁理士 伊藤 公一
(31)優先権主張番号	18/335,450	(74)代理人	100123593 弁理士 関根 宣夫
(32)優先日	令和5年6月15日(2023.6.15)	(74)代理人	100133835 弁理士 河野 努
(33)優先権主張国・地域又は機関	米国(US)		

最終頁に続く

(54)【発明の名称】 機械学習評価パイプラインについてのシステム及び方法

(57)【特許請求の範囲】

【請求項1】

機械学習(ML)モデルを評価する方法であって、
前記方法は、
ストレージ層から取得される少なくとも1つの要件を要件管理層によって受信することと、

前記少なくとも1つの要件を前記要件管理層によって解釈することと、
前記解釈された要件に基づいてML評価プロセスを行うための命令を前記要件管理層によって実行層に送信することであって、前記実行層は、前記ML評価プロセスを完了すると、前記ML評価プロセスの結果を有する出力信号を送信する、ということと、
を含む、方法。

10

【請求項2】

前記少なくとも1つの要件は、要件コード(RaC)ファイルの形態である、請求項1に記載の方法。

【請求項3】

前記ストレージ層は、ユーザが前記少なくとも1つの要件を編集することを可能にするように構成された第1のユーザインターフェースと通信する、請求項1又は2に記載の方法。

【請求項4】

前記実行層は、前記ML評価プロセスの前記結果を表示するように構成された第2のコ

20

ーザインターフェースに前記出力信号を送信するように構成されている、請求項 3 に記載の方法。

【請求項 5】

前記実行層は、推論コンポーネントと、ユニットテストコンポーネントと、を備え、ML 評価プロセスを行うための前記命令を受信すると、前記推論コンポーネントは、テストデータを受信して、前記テストデータ、及び前記 ML モデルから出力を取得するための命令に基づいて推論プロセスを行うように構成されており、前記ユニットテストコンポーネントは、前記 ML モデルからの前記出力、及び指標を取得するための前記命令に基づいて評価プロセスを行うように構成されている、請求項 4 に記載の方法。

【請求項 6】

前記 ML モデルからの前記出力は、推論ログを備え、前記推論プロセスを完了すると、前記推論コンポーネントは、前記推論ログを前記ユニットテストコンポーネントに転送するように構成されており、前記評価プロセスは、前記推論ログからの指標を評価することを含み、前記評価された指標は、前記第 2 のユーザインターフェースに表示される、請求項 5 に記載の方法。

【請求項 7】

前記方法は、前記ストレージ層において少なくとも 1 つの要件を追加又は更新するための命令を前記要件管理層によって受信することを更に含み、前記出力信号は、前記 ML 評価プロセスを完了すると、前記評価された指標に基づいて前記実行層から送信される、請求項 6 に記載の方法。

【請求項 8】

機械学習 (ML) モデルを評価する装置であって、
前記装置は、
コンピュータ実行可能命令を記憶する少なくとも 1 つのメモリと、
少なくとも 1 つのプロセッサと、
を備え、
前記少なくとも 1 つのプロセッサは、
ストレージ層から取得される少なくとも 1 つの要件を要件管理層によって受信することと、

前記少なくとも 1 つの要件を前記要件管理層によって解釈することと、

前記解釈された要件に基づいて ML 評価プロセスを行うための命令を前記要件管理層によって実行層に送信することであって、前記実行層は、前記 ML 評価プロセスを完了すると、前記 ML 評価プロセスの結果を有する出力信号を送信する、ということと、
を行うように、前記コンピュータ実行可能命令を実行するように構成されている、装置。

【請求項 9】

前記少なくとも 1 つの要件は、要件コード (R a C) ファイルの形態である、請求項 8 に記載の装置。

【請求項 10】

前記ストレージ層は、ユーザが前記少なくとも 1 つの要件を編集することを可能にするように構成された第 1 のユーザインターフェースと通信する、請求項 8 又は 9 に記載の装置。

【請求項 11】

前記実行層は、前記 ML 評価プロセスの前記結果を表示するように構成された第 2 のユーザインターフェースに前記出力信号を送信するように構成されている、請求項 10 に記載の装置。

【請求項 12】

前記実行層は、推論コンポーネントと、ユニットテストコンポーネントと、を備え、ML 評価プロセスを行うための前記命令を受信すると、前記推論コンポーネントは、テストデータを受信して、前記テストデータ、及び前記 ML モデルから出力を取得するための命令に基づいて推論プロセスを行うように構成されており、前記ユニットテストコンポーネ

10

20

30

40

50

ントは、前記MLモデルからの前記出力、及び指標を取得するための前記命令に基づいて評価プロセスを行うように構成されている、請求項11に記載の装置。

【請求項13】

前記MLモデルからの前記出力は、推論ログを備え、前記推論プロセスを完了すると、前記推論コンポーネントは、前記推論ログを前記ユニットテストコンポーネントに転送するように構成されており、前記評価プロセスは、前記推論ログからの指標を評価することを含み、前記評価された指標は、前記第2のユーザインターフェースに表示される、請求項12に記載の装置。

【請求項14】

前記プロセッサは更に、前記ストレージ層において少なくとも1つの要件を追加又は更新するための命令を前記要件管理層によって受信するように、前記コンピュータ実行可能命令を実行するように構成されており、前記出力信号は、前記ML評価プロセスを完了すると、前記評価された指標に基づいて前記実行層から送信される、請求項13に記載の装置。

10

【請求項15】

少なくとも1つのプロセッサによって実行可能な命令が記録されて前記プロセッサに方法を行わせる非一時的コンピュータ可読記録媒体であって、

前記方法は、

ストレージ層から取得される少なくとも1つの要件を要件管理層によって受信することと、

20

前記少なくとも1つの要件を前記要件管理層によって解釈することと、

前記解釈された要件に基づいてML評価プロセスを行うための命令を前記要件管理層によって実行層に送信することであって、前記実行層は、前記ML評価プロセスを完了すると、前記ML評価プロセスの結果を有する出力信号を送信する、ということと、を含む、非一時的コンピュータ可読記録媒体。

【請求項16】

前記少なくとも1つの要件は、要件コード(RaC)ファイルの形態である、請求項15に記載の非一時的コンピュータ可読記録媒体。

【請求項17】

前記ストレージ層は、ユーザが前記少なくとも1つの要件を編集することを可能にするように構成された第1のユーザインターフェースと通信する、請求項15又は16に記載の非一時的コンピュータ可読記録媒体。

30

【請求項18】

前記実行層は、前記ML評価プロセスの前記結果を表示するように構成された第2のユーザインターフェースに前記出力信号を送信するように構成されている、請求項17に記載の非一時的コンピュータ可読記録媒体。

【請求項19】

前記実行層は、推論コンポーネントと、ユニットテストコンポーネントと、を備え、ML評価プロセスを行うための前記命令を受信すると、前記推論コンポーネントは、テストデータを受信して、前記テストデータ、及びMLモデルから出力を取得するための命令に基づいて推論プロセスを行うように構成されており、前記ユニットテストコンポーネントは、前記MLモデルからの前記出力、及び指標を取得するための前記命令に基づいて評価プロセスを行うように構成されている、請求項18に記載の非一時的コンピュータ可読記録媒体。

40

【請求項20】

前記MLモデルからの前記出力は、推論ログを備え、前記推論プロセスを完了すると、前記推論コンポーネントは、前記推論ログを前記ユニットテストコンポーネントに転送するように構成されており、前記評価プロセスは、前記推論ログからの指標を評価することを含み、前記評価された指標は、前記第2のユーザインターフェースに表示される、請求項19に記載の非一時的コンピュータ可読記録媒体。

50

【発明の詳細な説明】**【技術分野】****【0001】**

本開示の例示的な実施形態と整合するシステム及び方法は、機械学習モデルを評価するパイプラインを提供することに関する。

【背景技術】**【0002】**

機械学習（ML）モデルは、様々なタスクを自動化するために使用され得る。MLモデルを開発する際、開発者は、MLモデルによって満たされる必要がある特定の基準又はパラメータを有し得る。例えば、MLモデルがセーフティクリティカルなタスクを自動化することを意図している場合、MLモデルは、特定の信頼性評価を達成する必要がある。したがって、MLモデルが規格を満たし得ることを保証するために、開発者は、ML評価を行うことによってMLモデルをテスト及び評価する必要がある。 10

【0003】

関連技術では、モデルが良いかどうかを自動的に決定するために単純な指標が使用され得る。例えば、平均適合率の平均（mAP）は、MLモデルが良いか、悪いか、又は優れているかどうかを示す単純な数として使用され得る。

【0004】

従来技術のシステム及び方法によって使用されるMLモデルを評価するプロセスは、限定的且つ低速であり得る。特に、通常、ML評価プロセスで使用される各コンポーネントは別々である。例えば、テストデータを取得する手段、MLモデル自体、ML評価テストユニットは全て、別々のシステム内にあり得る。更に、関連技術では、評価プロセスでの各ステップは、人間のユーザがML評価の結果を解釈して、より最適なMLモデルを取得するために評価プロセスをどのように繰り返すかを考慮することを必要とし得る。 20

【0005】

したがって、ML評価を行うために、より合理化及び自動化された方法が必要である。

【発明の概要】**【0006】**

1つ以上の例示的な実施形態によれば、機械学習（ML）モデルを評価する装置及び方法が提供される。特に、例示的な実施形態に係る装置及び方法は、ストレージ層から取得される少なくとも1つの要件を要件管理層において受信し、要件を要件管理層において解釈し、解釈されたテストパラメータに基づいてML評価プロセスを行うための命令を要件管理層によって実行層に送信する。ML評価プロセスを完了すると、実行により、ML評価プロセスの結果を有する出力信号を送信し得る。出力信号に基づいて、結果の情報は、ユーザインターフェース（UI）に表示され得る。したがって、ML評価プロセスを構成及び実行する全プロセスを、単一のパイプラインに合理化/カプセル化することができ、任意選択的に、パイプラインの層に接続されたユーザインターフェースを用いてユーザに提示できるため、評価プロセスの自動化を改善できる。 30

【0007】

実施形態によれば、MLモデルを評価する方法が提供され得る。方法は、ストレージ層から取得される少なくとも1つの要件を要件管理層によって受信することと、少なくとも1つの要件を要件管理層によって解釈することと、解釈されたテストパラメータに基づいてML評価プロセスを行うための命令を要件管理層によって実行層に送信することと、実行層は、ML評価プロセスを完了すると、ML評価プロセスの結果を有する出力信号を送信する、ということと、を含み得る。 40

【0008】

少なくとも1つのテストパラメータは、要件コード（R a C）ファイルの形態であり得る。

【0009】

ストレージ層は、ユーザが少なくとも1つのテストパラメータを編集することを可能に 50

するように構成された第1のユーザインターフェースと通信し得る。

【0010】

実行層は、ML評価プロセスの結果を表示するように構成された第2のユーザインターフェースに出力信号を送信するように構成され得る。

【0011】

第1のユーザインターフェース及び第2のユーザインターフェースは、同時に表示され得る。

【0012】

実行層は、推論コンポーネントと、ユニットテストコンポーネントと、を含み得、ML評価プロセスを行うための命令を受信すると、推論コンポーネントは、テストデータを受信して、テストデータ、及びMLモデルから出力を取得するための命令に基づいて推論プロセスを行うように構成されており、ユニットテストコンポーネントは、MLモデルからの出力、及び指標を取得するための命令に基づいて評価プロセスを行うように構成されている。

10

【0013】

推論コンポーネントは、テストデータストレージ層からテストデータを受信するように構成され得る。

【0014】

実施形態によれば、機械学習(ML)モデルを評価する装置が提供され得る。装置は、コンピュータ実行可能命令を記憶する少なくとも1つのメモリと、少なくとも1つのプロセッサと、を含み得、少なくとも1つのプロセッサは、ストレージ層から取得される少なくとも1つの要件を要件管理層によって受信することと、少なくとも1つの要件を要件管理層によって解釈することと、解釈された要件に基づいてML評価プロセスを行うための命令を要件管理層によって実行層に送信することと、実行層は、ML評価プロセスを完了すると、ML評価プロセスの結果を有する出力信号を送信する、ということと、を行うように、コンピュータ実行可能命令を実行するように構成されている。

20

【0015】

追加の態様は、部分的に以下の説明で記載され、部分的に説明から明らかになるか、又は本開示の提示される実施形態の実施によって実現され得る。

【図面の簡単な説明】

30

【0016】

本開示の特定の好ましい実施形態の特徴、態様、及び利点は、添付図面を参照して以下に記載され、当該添付図面では、同様の参照番号は同様の要素を示す。

【図1】図1は、例示的な実施形態に係るデバイスの例示的なコンポーネントの図である。

【図2】図2は、1つ以上の例示的な実施形態に係る層及びコンポーネントを示すシステムアーキテクチャ図である。

【図3】図3は、1つ以上の例示的な実施形態に係るMLモデルを評価する方法を示すフローチャート図である。

【図4】図4は、1つ以上の例示的な実施形態に係る、推論コンポーネント及びユニットテストコンポーネントを使用してMLモデルを評価する方法を示すフローチャート図である。

40

【図5】図5は、1つ以上の例示的な実施形態に係る、要件を追加又は更新することを含むMLモデルを評価する方法を示すフローチャート図である。

【発明を実施するための形態】

【0017】

例示的な実施形態の以下の詳細な説明は、添付図面を参照する。本開示は、例示及び説明を提供するが、網羅的であることを意図したものでもなく、1つ以上の例示的な実施形態を、開示される正確な形態に限定することを意図したものでもない。修正及び変形は、本開示を鑑みて可能であるか、又は1つ以上の例示的な実施形態の実施から習得され得る。更に、例示的な一実施形態の1つ以上の特徴又はコンポーネントは、例示的な別の実施

50

形態（若しくは例示的な別の実施形態の1つ以上の特徴）に組み込まれ得るか、又はそれと組み合わせられ得る。更に、本明細書で提供される動作のフローチャート及び説明において、1つ以上の動作が省略されてもよく、1つ以上の動作が追加されてもよく、1つ以上の動作が（少なくとも部分的に）同時に行われてもよく、1つ以上の動作の順序が切り替えられてもよいことが理解される。

【0018】

本明細書に記載されるシステム及び/若しくは方法並びに/又は非一時的コンピュータ可読記憶媒体の例示的な実施形態は、ハードウェア、ファームウェア、又はハードウェア及びソフトウェアの組合せの様々な形態で実装され得ることが明らかであろう。当該システム及び/又は方法を実装するために使用される実際の専用の制御ハードウェア又はソフトウェアコードは、1つ以上の例示的な実施形態の限定ではない。したがって、システム及び/若しくは方法並びに/又は非一時的コンピュータ可読記憶媒体の動作及び挙動は、特定のソフトウェアコードを参照することなく本明細書に記載される。ソフトウェア及びハードウェアは、本明細書の説明に基づいてシステム及び/又は方法を実装するように設計され得ることが理解される。

10

【0019】

特徴の特定の組合せが特許請求の範囲で列挙され及び/又は本明細書に開示されていても、当該組合せは、可能性のある例示的な実施形態の開示を限定することを意図したものではない。実際、当該特徴の多くは、具体的に特許請求の範囲で列挙されていない方法及び/又は本明細書に開示されていない方法で組み合わせられ得る。以下に列挙される各従属請求項は、1つの請求項のみに直接的に従属し得るが、可能性のある例示的な実施形態の開示は、請求項のセットにおける全ての他の請求項と組み合わせられた各従属請求項を含む。

20

【0020】

本明細書で使用される要素、行為、又は命令は、特に明示的に記載されていない限り、重要又は必須であると解釈されるべきではない。また、本明細書で使用される冠詞「a」及び「an」は、1つ以上の事項を含むことを意図したものであり、「1つ以上」と交換可能に使用され得る。1つの事項のみを意図したものである場合、「1つ」という用語又は同様の用語が使用される。また、本明細書で使用される用語「有する(has)」、「有する(have)」、「有している(having)」、「含む(include)」、「含んでいる(including)」、又は同種のもは、オープンエンドの用語であることを意図したものである。更に、「~に基づいて」というフレーズは、特に明示的に述べられていない限り、「~に少なくとも部分的に基づいて」を意味することを意図したものである。更に、「[A]及び[B]のうち少なくとも一方」又は「[A]又は[B]のうち少なくとも一方」などの表現は、Aのみ、Bのみ、又はA及びBの両方を含むものとして理解されるべきである。

30

【0021】

図1は、機械学習(ML)評価デバイス100の例示的なコンポーネントの図である。図1に示されるように、ML評価デバイス100は、バス110と、プロセッサ120と、メモリ130と、ストレージコンポーネント140と、入力コンポーネント150と、出力コンポーネント160と、通信インターフェース170と、を含み得る。

40

【0022】

バス110は、ML評価デバイス100のコンポーネント間の通信を可能にするコンポーネントを含む。プロセッサ120は、ハードウェア、ファームウェア、又はハードウェア及びソフトウェアの組合せで実装され得る。プロセッサ120は、中央処理装置(CPU)、画像処理装置(GPU)、加速処理装置(APU)、マイクロプロセッサ、マイクロコントローラ、デジタルシグナルプロセッサ(DSP)、フィールドプログラマブルゲートアレイ(FPGA)、特定用途向け集積回路(ASIC)、又は別のタイプの処理コンポーネントであり得る。1つ以上の例示的な実施形態では、プロセッサ120は、機能を行うようにプログラム可能な1つ以上のプロセッサを含む。メモリ130は、プロセッサ220による使用のために情報及び/又は命令を記憶するランダムアクセスメモリ(R

50

AM)、リードオンリーメモリ(ROM)、並びに/又は別のタイプの動的若しくは静的ストレージデバイス(例えば、フラッシュメモリ、磁気メモリ、及び/若しくは光メモリ)を含む。

【0023】

ストレージコンポーネント140は、ML評価デバイス100の動作及び使用に関連する情報及び/又はソフトウェアを記憶する。例えば、ストレージコンポーネント140は、対応するドライブと共に、ハードディスク(例えば、磁気ディスク、光ディスク、光磁気ディスク、及び/若しくはソリッドステートディスク)、コンパクトディスク(CD)、デジタルバーサタイルディスク(DVD)、フロッピーディスク、カートリッジ、磁気テープ、並びに/又は別のタイプの非一時的コンピュータ可読媒体を含み得る。入力コンポーネント150は、ユーザ入力などを介して情報をML評価デバイス100が受信することを可能にするコンポーネント(例えば、タッチスクリーンディスプレイ、キーボード、キーパッド、マウス、ボタン、スイッチ、及び/又はマイク)を含む。追加的に又は代替的に、入力コンポーネント150は、情報を検知するセンサ(例えば、グローバルポジショニングシステム(GPS)コンポーネント、加速度計、ジャイロスコープ、及び/又はアクチュエータ)を含み得る。出力コンポーネント160は、ML評価デバイス100からの出力情報を提供するコンポーネント(例えば、ディスプレイ、スピーカ、及び/又は1つ以上の発光ダイオード(LED))を含む。

10

【0024】

通信インターフェース170は、有線接続、無線接続、又は有線及び無線接続の組合せなどを介してML評価デバイス100が他のデバイスと通信することを可能にする送受信機状のコンポーネント(例えば、送受信機及び/又は別々の受信機及び送信機)を含む。通信インターフェース170は、ML評価デバイス100が別のデバイスから情報を受信すること及び/又は情報を別のデバイスに提供することを可能にし得る。例えば、通信インターフェース170は、イーサネットインターフェース、光インターフェース、同軸インターフェース、赤外線インターフェース、無線周波数(RF)インターフェース、ユニバーサルシリアルバス(USB)インターフェース、Wi-Fiインターフェース、セルラネットワークインターフェース、又は同種のものを含み得るが、これらに限定されない。

20

【0025】

ML評価デバイス100は、本明細書に記載される1つ以上の例示的なプロセスを行い得る。1つ以上の例示的な実施形態によれば、ML評価デバイス100は、メモリ130及び/又はストレージコンポーネント140などの非一時的コンピュータ可読媒体によって記憶されたソフトウェア命令をプロセッサ120が実行したことに応じて当該プロセスを行い得る。コンピュータ可読媒体は、本明細書において非一時的メモリデバイスとして定められる。メモリデバイスは、単一の物理ストレージデバイス内のメモリ空間、又は複数の物理ストレージデバイスに及ぶメモリ空間を含む。

30

【0026】

ソフトウェア命令は、通信インターフェース170を介して別のコンピュータ可読媒体から又は別のデバイスからメモリ130及び/又はストレージコンポーネント140に読み込まれ得る。メモリ130及び/又はストレージコンポーネント140に記憶されたソフトウェア命令は、実行されると、本明細書に記載される1つ以上のプロセスをプロセッサ120に行わせ得る。

40

【0027】

追加的に又は代替的に、ハードワイヤード回路は、本明細書に記載される1つ以上のプロセスを行うために、ソフトウェア命令の代わりに、又はソフトウェア命令と組み合わされて使用され得る。したがって、本明細書に記載される1つ以上の例示的な実施形態は、ハードウェア回路及びソフトウェアのどの特定の組合せにも限定されない。

【0028】

図1に示されるコンポーネントの数及び配置は、一例として提供される。実際、ML評価デバイス100は、図1に示されるものと比べて、追加のコンポーネント、より少ない

50

コンポーネント、異なるコンポーネント、又は異なって配置されたコンポーネントを含み得る。追加的に又は代替的に、ML評価デバイス100のコンポーネント(例えば、1つ以上のコンポーネント)のセットは、ML評価デバイス100のコンポーネントの別のセットによって行われるものとして記載される1つ以上の機能を行い得る。

【0029】

図2は、1つ以上の例示的な実施形態に係る層及びコンポーネントを示すシステムアーキテクチャ図である。層(例えば、実行層200、要件管理層210、及びストレージ層220)は、組み合わされて機械学習(ML)評価パイプラインと見なされ得る。

【0030】

実施形態によれば、実行層200が提供され得る。実行層200は、推論コンポーネント201と、ユニットテストコンポーネント202と、を含み得る。実行層200は、MLモデル240に対するML評価プロセスの実行を担い得る。実行層200は、ML評価を実行する任意の適切な手段を実装し得ることを理解されたい。実行層200は、要件管理層210から命令を受信するように構成され得る。実行層200はまた、評価の結果を出力するために表示結果UI260とやり取りし得る。特に、実行層200は、評価の完了時に評価の結果を含む出力信号を送信し得、一部の実施形態によれば、表示結果UI260は、出力信号を受信して評価の結果を表示し得る。表示結果UI260は、任意の適切なUIを用いて実装され得る。例えば、表示結果UI260は、評価テストの結果を列挙するために表形式を有してもよく、表示結果UI260は、評価テストが合格であるか不合格であるかに基づいてテスト結果の特定の部分を強調してもよい。それにも関わらず、表示UI260は、当業者によって決定されるように特定の实装態様に依存し得ることを理解されたい。

10

20

【0031】

実行層200はまた、要件管理層210を介して、又はテストデータストレージ222から直接的に、テストデータストレージ222からテストデータを受信し得る。

【0032】

一部の実施形態によれば、推論コンポーネント201は、テストデータストレージ222からのデータの受信、及びMLモデル240を使用して出力を計算するためのMLモデル240へのデータの入力(すなわち、推論)を担い得る。

【0033】

一部の実施形態によれば、ユニットテストコンポーネント202は、MLモデル240に対するユニットテストの実行を担い得る。具体的には、ユニットテストは、MLモデル240に対して推論コンポーネント201を使用して行われる推論プロセス中に取得されるMLモデル240からの出力の結果を評価し得る。MLモデル240の性能に関する指標は、ユニットテストコンポーネント202を使用して取得され得る。

30

【0034】

要件管理層210が提供され得る。要件管理層210は、ストレージ層に記憶されたファイル(例えば、要件コード(RAC)ファイル1、2、...、N(221-1、221-2、221-N、...))又はテストデータストレージ222からのテストデータの解釈を担い得る。RACファイルの解釈に基づいて、要件管理層210は、ML評価を実行するように実行層200に命令を送信し得る。

40

【0035】

ストレージ層220は、実行層200及び要件管理層210によって使用される全てのデータを含み得る。ストレージ層220は、任意の適切なストレージ手段(例えば、データベース、クラウドストレージなど)によって実装され得る。ストレージ層220は、任意の数のRACファイル221-1、221-2、221-Nと、テストデータストレージ222と、を含み得る。各RACファイル及びテストデータストレージは、同じ記憶媒体又は異なる記憶媒体に記憶され得ることを理解されたい。

【0036】

要件コード(RAC)ファイル221-1、221-2、221-Nは、MLモデル2

50

40の期待される挙動の記憶を担い得る。期待される挙動は、(テストパラメータを含み得る)要件を備え得る。特に、RACファイルは、例えば、達成される必要がある特定の指標基準、達成される必要がある規格、テストのタイプなどに基づいて、テストされる必要があるMLモデルの性能指標を指定する人間可読形式であり得る。RACファイルは、実行層200においてML評価プロセス中に使用される要件、合格基準、テスト目標、テストデータのファイルパス、テスト条件を含み得る。したがって、このような要件、合格基準、テスト目標、テストデータパス、及びテスト条件は、どのようにMLモデル240のMLテスト及び評価が実行層200によって実行されるべきかを決定するために、要件管理層210によって容易に解釈され得る。

【0037】

一部の実施形態によれば、RACファイル221-1、221-2、221-Nは、編集プロンプトUI250によって編集可能であり得る。特に、編集プロンプトUI250は、ソースファイルを編集するために使用される汎用コマンドラインインターフェースであり得るか、又は編集プロンプトUI250は、ユーザが所定の構成をドラッグアンドドロップ又は選択することを可能にするようにやり取り可能な要素を有するグラフィカルユーザインターフェース(GUI)であり得る。それにも関わらず、任意の適切なユーザインターフェースが編集プロンプトUI250に対して実装され得ることを理解されたい。

【0038】

テストデータストレージ222は、推論コンポーネント201によって使用されてMLモデル240に入力されるデータを含み得る。一部の実施形態によれば、テストデータストレージ222は、RACファイルと別々に記憶され得る。テストデータについてのデータファイルの特定の形式は、任意の適切な形式であり得る。

【0039】

編集プロンプトUI250及び表示結果UI260は、任意の適切な環境、例えば、ウェブインターフェースにおいて、又はローカルユーザデバイス上でのみ実装され得ることを理解されたい。一部の実施形態によれば、編集プロンプトUI250及び表示結果UI260は、別々に又は同時に表示され得ることも理解されたい。

【0040】

実施形態によれば、(編集プロンプトUI250及び表示結果UI260の特徴を含み得る)UIが提供されてもよく、これは、評価プロセスを合理化するための更なる特徴を含み得る。例えば、ユーザインターフェースは、ストレージ層220からのどのセットの要件が使用されるべきか、及びその順序付け、例えば、第1のセット及び第2のセットの要件が順次評価されるべきか又は並列で評価されるべきかをユーザが選択することを可能にするためのインターフェースと共に要件管理層210と通信し得る。上記ユーザインターフェースはまた、(例えば、テストデータストレージ222から)ユーザがテストデータセットを選択することを可能にし得る。ユーザは、異なる候補MLモデルにわたって動作して、どの候補MLモデルが最高の指標を有するかを自動的に示すように同じセットの要件を構成でき得ることも想定される。

【0041】

図3は、1つ以上の例示的な実施形態に係るMLモデルを評価する方法300を示すフローチャート図である。

【0042】

図3を参照して、動作S310で、テストパラメータを含む要件は、ストレージ層220から要件管理層210によって受信される。一部の実施形態によれば、(テストパラメータなどの)要件は、RACファイル221-1、221-2、221-Nのうちの1つ以上であり得る。

【0043】

ストレージ層220から要件を受信した後、動作S320で、要件は、MLモデル240に対してML評価を行うように実行層200にどのように命令するかを決定するために要件管理層210によって解釈され得る。これは、要件、合格基準、テスト目標、テスト

10

20

30

40

50

データパス、信頼閾値などのモデル調整パラメータ、要件によるテスト条件を解釈することと、実行層 200 を使用して行われるべき適切なテスト及びテストパラメータを決定することと、を含み得る。したがって、解釈される要件は、要件管理層 210 によって取得され得る。

【0044】

動作 S330 で、要件管理層 210 は、ML モデル 240 に対して ML 評価を実行層 200 に行わせるために、動作 S320 から解釈されたテストパラメータに基づいて実行層 200 に命令を送信する。一部の実施形態によれば、実行層 200 が ML 評価の実行を終了すると、実行層 200 は、評価の結果を含む出力信号を送信し得る。一部の実施形態によれば、この出力信号は、表示結果 UI 260 に送信され得る。

10

【0045】

図 4 は、1 つ以上の例示的な実施形態に係る、推論コンポーネント及びユニットテストコンポーネントを使用して ML モデルを評価する方法 400 を示すフローチャート図である。図 2 に示されるものと同様の層及びコンポーネントは、方法 400 を実装するために使用され得、図 3 に示されるような方法 300 と同様のステップの完全な説明は、読みやすくするために省略され得る。

【0046】

図 4 を参照して、動作 S410 で、ML 評価を行うための命令は、要件管理層 210 から送信されて、ML モデル 240 に対して ML 評価を実行層 200 に行わせるために実行層 200 によって受信され得る。動作 S410 で受信される命令は、図 3 で上述したような動作 S330 で送信されるものと同様であり得る。

20

【0047】

動作 S420 で、実行層 200 は、テストデータ（例えば、テストデータストレージ 222 から取得されるテストデータ）と共に、ML モデル 240 を使用して推論動作を実行するように推論コンポーネント 201 に命令を送信し得る。この動作の出力は、推論ログを含み得、当該推論ログは、テストデータを使用した ML モデル 240 に対する推論動作の実行からのデータを含み得る。

【0048】

動作 S430 で、実行層 200 は、推論コンポーネント 201 からユニットテストコンポーネント 202 に推論ログを送信し得る。

30

【0049】

動作 S440 で、実行層 200 は、（例えば、要件によるテスト基準に関して指標を計算及び比較することによって）指標を評価するようにユニットテストコンポーネント 202 に命令を送信し得る。

【0050】

動作 S450 で、実行層 200 は、（比較結果を含み得る）評価された指標を出力し得る。一部の実施形態によれば、これは、表示結果 UI 260 などの UI に指標及び比較結果を送信することを含み得る。

【0051】

図 5 は、1 つ以上の例示的な実施形態に係る、要件を追加又は更新することを含む ML モデルを評価する方法を示すフローチャート図である。図 2 に示されるものと同様の層及びコンポーネントは、方法 500 を実装するために使用され得、図 3 に示されるような方法 300 及び上記の図 4 に示されるような方法 400 と同様のステップの完全な説明は、読みやすくするために省略され得る。

40

【0052】

図 5 を参照して、動作 S510 で、テストパラメータを含む要件は、ストレージ層 220 から要件管理層 210 によって受信される。これは、図 3 で上述した動作 S310 と同様であり得る。

【0053】

ストレージ層 220 から要件を受信した後、動作 S520 で、要件は、ML モデル 24

50

0 に対して M L 評価を行うように実行層 2 0 0 にどのように命令するかを決定するために要件管理層 2 1 0 によって解釈され得る。これは、図 3 で上述した動作 S 3 2 0 と同様であり得る。

【 0 0 5 4 】

動作 S 5 3 0 で、要件管理層 2 1 0 は、M L モデル 2 4 0 に対して M L 評価を実行層 2 0 0 に行わせるために、動作 S 5 2 0 から解釈されたテストパラメータに基づいて実行層 2 0 0 に命令を送信する。これは、図 3 で上述した動作 S 3 3 0 と同様であり得る。

【 0 0 5 5 】

動作 S 5 4 0 で、M L 評価を行うための命令は、M L モデル 2 4 0 に対して M L 評価を実行層 2 0 0 に行わせるために実行層 2 0 0 によって受信され得る。これは、図 4 で上述した動作 S 4 1 0 と同様であり得る。

【 0 0 5 6 】

動作 S 5 5 0 で、命令実行層 2 0 0 は、(例えば、要件によるテスト基準に関して指標を計算及び比較することによって) 指標を評価するようにユニットテストコンポーネント 2 0 2 に命令を送信し得る。これは、図 4 で上述した動作 S 4 4 0 と同様であり得る。明示的に示されていないが、一部の実施形態によれば、S 4 2 0 ~ S 4 3 0 と同様の動作も動作 S 5 5 0 の前に含まれ得ることを理解されたい。

【 0 0 5 7 】

動作 S 5 6 0 で、実行層 2 0 0 は、(比較結果を含み得る) 評価された指標を出力して、表示結果 U I 2 6 0 などの U I に指標及び比較結果を送信し得る。これは、図 4 で上述した動作 S 4 5 0 と同様であり得る。

【 0 0 5 8 】

動作 S 5 7 0 で、実行層 2 0 0 は、(例えば、新しい R A C ファイルをストレージ層 2 2 0 に追加することによって) 要件をストレージ層 2 2 0 に追加するように、又は(例えば、ストレージ層 2 2 0 内の既存の R A C ファイルを編集することによって) 既存の要件を更新するように要件管理層 2 1 0 に命令を送信し得る。命令は、M L モデル 2 4 0 を含むシステムの全体的な性能を最適化するための信頼閾値及びテストパラメータなどであるが、これらに限定されない M L モデル調整パラメータを更新するために送信され得る。

【 0 0 5 9 】

動作 S 5 8 0 で、要件管理層 2 1 0 は、動作 S 5 7 0 で送信された命令を受信し、それに従って、ストレージ層 2 2 0 において要件を追加又は更新する。その後、ステップは、S 5 1 0 から繰り返され得る。本実施形態によれば、まだ発見されていない場合がある要件及びパラメータを探索する反復プロセスが実装され得る。

【 0 0 6 0 】

上記実施形態に基づいて、M L 評価プロセスを構成及び実行する全プロセスを、単一のパイプラインに合理化/カプセル化することができ、任意選択的に、パイプラインの層に接続されたユーザインターフェースを用いてユーザに提示できるため、評価プロセスの自動化を改善できることが理解され得る。

【 0 0 6 1 】

以上の開示は、例示及び説明を提供するが、網羅的であることを意図したものでもなく、1 つ以上の例示的な実施形態を、開示される正確な形態に限定することを意図したものでもない。修正及び変形は、本開示を鑑みて可能であるか、又は 1 つ以上の例示的な実施形態の実施から習得され得る。

【 0 0 6 2 】

1 つ以上の例示的な実施形態は、統合の任意の可能性がある技術的詳細レベルにおけるシステム、方法、及び/又はコンピュータ可読媒体に関し得る。更に、上述のコンポーネントのうち 1 つ以上は、コンピュータ可読媒体において記憶され少なくとも 1 つのプロセッサによって実行可能な命令として実装され得る(及び/又は少なくとも 1 つのプロセッサを含み得る)。コンピュータ可読媒体は、プロセッサに動作を実行させるためにコンピュータ可読プログラム命令を有するコンピュータ可読非一時的記憶媒体(又は媒体(複

10

20

30

40

50

数)) を含み得る。

【 0 0 6 3 】

コンピュータ可読記憶媒体は、命令実行デバイスによって使用される命令を保持及び記憶し得る有形デバイスであり得る。コンピュータ可読記憶媒体は、例えば、電子ストレージデバイス、磁気ストレージデバイス、光ストレージデバイス、電磁ストレージデバイス、半導体ストレージデバイス、又は以上の任意の好適な組合せであり得るが、これらに限定されない。コンピュータ可読記憶媒体のより特定の例の非網羅的なリストは、以下、すなわち、ポータブルコンピュータディスク、ハードディスク、ランダムアクセスメモリ (RAM)、リードオンリーメモリ (ROM)、消去可能なプログラマブルリードオンリーメモリ (EPROM 又はフラッシュメモリ)、静的ランダムアクセスメモリ (SRAM)、ポータブルコンパクトディスクリードオンリーメモリ (CD-ROM)、デジタルバーサタイルディスク (DVD)、メモリスティック、フロッピーディスク、命令が記録されたパンチカード又は溝内の隆起構造などの機械的に符号化されたデバイス、及び以上の任意の好適な組合せを含む。本明細書で使用されるコンピュータ可読記憶媒体は、電波若しくは他の自由に伝播する電磁波、導波路若しくは他の送信媒体を通じて伝播する電磁波 (例えば、光ファイバケーブルを通過する光パルス)、又はワイヤを通じて送信される電気信号などの、一時的な信号自体であると解釈されるべきではない。

10

【 0 0 6 4 】

本明細書に記載されるコンピュータ可読プログラム命令は、コンピュータ可読記憶媒体からそれぞれの計算 / 処理デバイスにダウンロードされ得るか、又はネットワーク、例えば、インターネット、ローカルエリアネットワーク、ワイドエリアネットワーク、及び / 若しくは無線ネットワークを介して外部コンピュータ若しくは外部ストレージデバイスにダウンロードされ得る。ネットワークは、銅送信ケーブル、光送信ファイバ、無線送信、ルータ、ファイアウォール、スイッチ、ゲートウェイコンピュータ、及び / 又はエッジサーバを備え得る。各計算 / 処理デバイス内のネットワークアダプタカード又はネットワークインターフェースは、ネットワークからコンピュータ可読プログラム命令を受信し、それぞれの計算 / 処理デバイス内のコンピュータ可読記憶媒体に記憶するためにコンピュータ可読プログラム命令を転送する。

20

【 0 0 6 5 】

動作を実行するコンピュータ可読プログラムコード / 命令は、アセンブラ命令、命令セットアーキテクチャ (ISA) 命令、機械命令、機械依存命令、マイクロコード、ファームウェア命令、状態設定データ、集積回路用の構成データ、又は 1 つ以上のプログラミング言語の任意の組合せで書かれたソースコード若しくはオブジェクトコードであり得、当該プログラミング言語は、Smalltalk、C++、又は同種のものなどのオブジェクト指向プログラミング言語と、「C」プログラミング言語又は同様のプログラミング言語などの手続き型プログラミング言語と、を含む。コンピュータ可読プログラム命令は、ユーザのコンピュータ上で完全に実行され得るか、ユーザのコンピュータ上で部分的に実行され得るか、スタンドアロンソフトウェアパッケージとして実行され得るか、ユーザのコンピュータ上で部分的に且つリモートコンピュータ上で部分的に実行され得るか、又はリモートコンピュータ若しくはサーバ上で完全に実行され得る。後者のシナリオでは、リモートコンピュータは、ローカルエリアネットワーク (LAN) 若しくはワイドエリアネットワーク (WAN) を含む任意のタイプのネットワークを通じてユーザのコンピュータに接続され得るか、又は (例えば、インターネットサービスプロバイダを使用してインターネットを通じて) 外部コンピュータへの接続が行われ得る。1 つ以上の例示的な実施形態では、例えば、プログラマブルロジック回路、フィールドプログラマブルゲートアレイ (FPGA)、又はプログラマブルロジックアレイ (PLA) を含む電子回路は、態様又は動作を行うために、コンピュータ可読プログラム命令の状態情報を利用して電子回路をパーソナライズすることによってコンピュータ可読プログラム命令を実行し得る。

30

40

【 0 0 6 6 】

当該コンピュータ可読プログラム命令は、マシンを生成するために汎用コンピュータ、

50

専用コンピュータ、又は他のプログラマブルデータ処理装置のプロセッサに提供され得、その結果、コンピュータ又は他のプログラマブルデータ処理装置のプロセッサを介して実行される命令は、フローチャート及び/又はブロック図のブロック若しくはブロック（複数）において指定される機能/行為を実装する手段を生成する。当該コンピュータ可読プログラム命令はまた、コンピュータ、プログラマブルデータ処理装置、及び/又は他のデバイスに特定の方法で機能するように指示し得るコンピュータ可読記憶媒体に記憶され得、その結果、命令が内部に記憶されたコンピュータ可読記憶媒体は、フローチャート及び/又はブロック図のブロック若しくはブロック（複数）において指定される機能/行為の態様を実装する命令を含む製造物品を備える。

【0067】

コンピュータ可読プログラム命令はまた、コンピュータ、他のプログラマブルデータ処理装置、又は他のデバイス上にロードされて、コンピュータ、他のプログラマブル装置、又は他のデバイス上で一連の動作ステップを行わせて、コンピュータ実装プロセスを生成し得、その結果、コンピュータ、他のプログラマブル装置、又は他のデバイス上で実行される命令は、フローチャート及び/又はブロック図のブロック若しくはブロック（複数）において指定される機能/行為を実装する。

【0068】

図におけるフローチャート及びブロック図は、1つ以上の例示的な実施形態に係るシステム、方法、及びコンピュータ可読媒体の可能性のある例示的な実施形態のアーキテクチャ、機能、及び動作を示す。この点に関して、フローチャート又はブロック図における各ブロックは、指定された論理機能を実装する1つ以上の実行可能命令を備える命令のマイクロサービス、モジュール、セグメント、又は一部を表し得る。方法、コンピュータシステム、及びコンピュータ可読媒体は、図面で描写されるものと比べて、追加のブロック、より少ないブロック、異なるブロック、又は異なって配置されたブロックを含み得る。1つ以上の代替的な例示的な実施形態では、ブロックに記される機能は、図に記される順序と関係なく生じ得る。例えば、連続して示される2つのブロックは実際、同時に若しくは実質的に同時に実行され得るか、又はブロックは、関連する機能に応じて、逆の順序で実行されることもあり得る。ブロック図及び/又はフローチャート図の各ブロック、並びにブロック図及び/又はフローチャート図のブロックの組合せは、指定された機能若しくは行為を行うか又は専用のハードウェア及びコンピュータ命令の組合せを実行する専用のハードウェアベースのシステムによって実装され得ることに留意されたい。

【0069】

本明細書に記載されるシステム及び/又は方法は、ハードウェア、ファームウェア、又はハードウェア及びソフトウェアの組合せの様々な形態で実装され得ることが明らかであろう。当該システム及び/又は方法を実装するために使用される実際の専用の制御ハードウェア又はソフトウェアコードは、1つ以上の例示的な実施形態の限定ではない。したがって、システム及び/又は方法の動作及び挙動は、特定のソフトウェアコードを参照することなく本明細書に記載されており、ソフトウェア及びハードウェアは、本明細書の説明に基づいてシステム及び/又は方法を実装するように設計され得ることが理解される。

10

20

30

40

50

【 図面 】

【 図 1 】

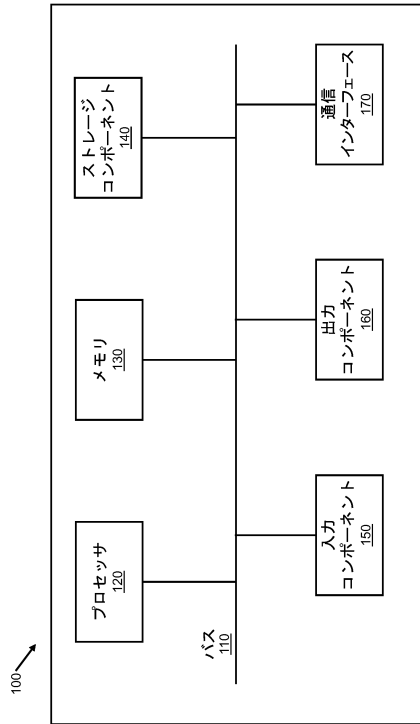
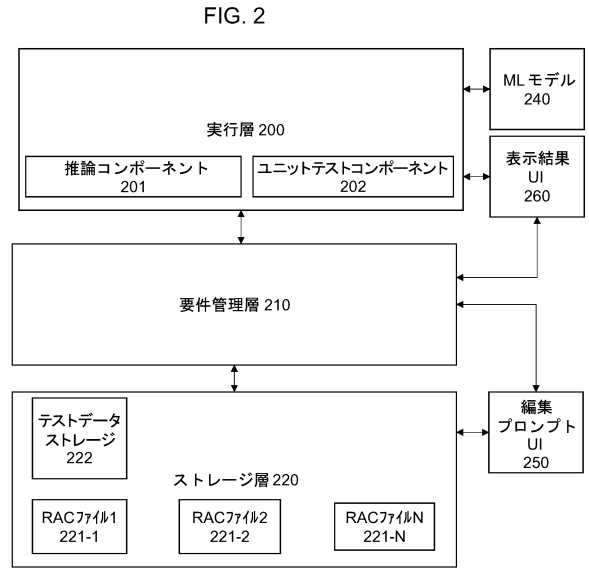


FIG. 1

【 図 2 】



10

20

【 図 3 】

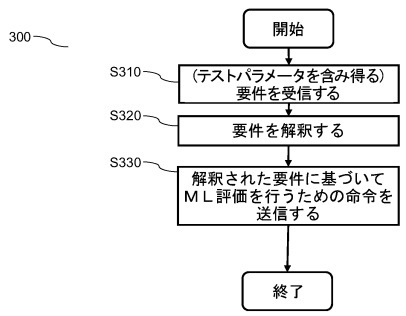


FIG. 3

【 図 4 】

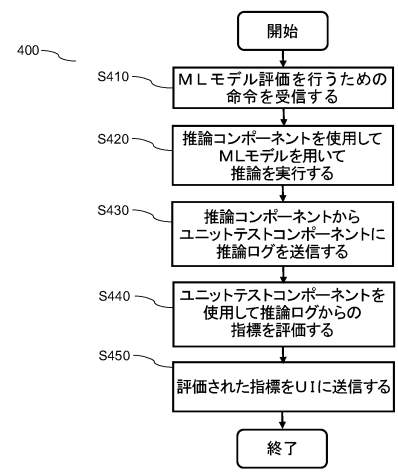


FIG. 4

30

40

【 図 5 】

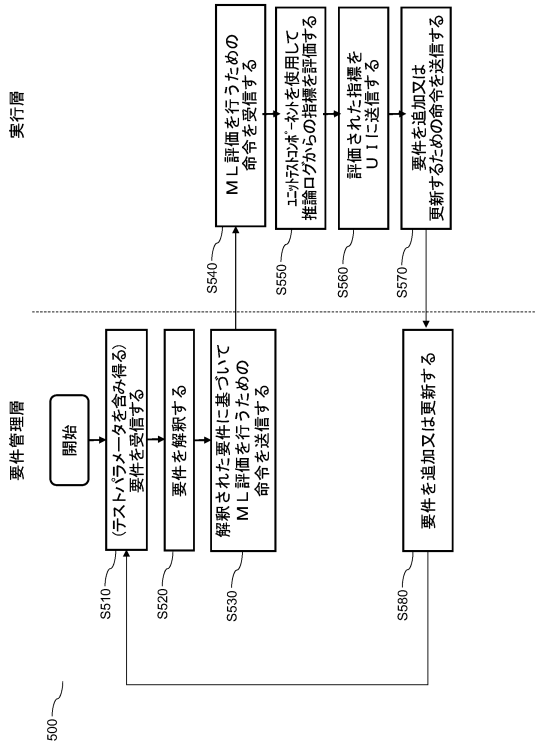


FIG. 5

10

20

30

40

50

フロントページの続き

- (74)代理人 100120499
弁理士 平山 淳
- (72)発明者 橋本 大輔
東京都中央区日本橋室町三丁目2番1号 ウーブン・バイ・トヨタ株式会社内
- (72)発明者 中津川 英治
東京都中央区日本橋室町三丁目2番1号 ウーブン・バイ・トヨタ株式会社内
- (72)発明者 小出 粹玄
東京都中央区日本橋室町三丁目2番1号 ウーブン・バイ・トヨタ株式会社内
- 審査官 大倉 峻吾
- (56)参考文献 特開2021-018508(JP,A)
特開2019-096285(JP,A)
米国特許出願公開第2022/0300754(US,A1)
米国特許出願公開第2022/0194423(US,A1)
- (58)調査した分野 (Int.Cl., DB名)
G06N 3/00-99/00
G06F 18/00-18/40
G06F 11/36