

(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号
特許第7704331号
(P7704331)

(45)発行日 令和7年7月8日(2025.7.8)

(24)登録日 令和7年6月30日(2025.6.30)

(51)国際特許分類		F I	
G 0 6 N	20/00 (2019.01)	G 0 6 N	20/00 1 3 0
G 0 6 N	3/0475(2023.01)	G 0 6 N	3/0475
G 0 6 N	3/0895(2023.01)	G 0 6 N	3/0895
G 0 6 F	16/90 (2019.01)	G 0 6 F	16/90 1 0 0

請求項の数 19 外国語出願 (全24頁)

(21)出願番号	特願2024-98979(P2024-98979)	(73)特許権者	514322098
(22)出願日	令和6年6月19日(2024.6.19)		ベイジン バイドゥ ネットコム サイエ ンス テクノロジー カンパニー リミテ ッド
(65)公開番号	特開2024-163990(P2024-163990 A)		Beijing Baidu Netco m Science Technolog y Co., Ltd.
(43)公開日	令和6年11月26日(2024.11.26)		中華人民共和国 ベキン 100085, ハイディアン ディストリクト, シャン ディ テンス ストリート, 10番, バ イドゥ キャンパス 2階
審査請求日	令和6年8月26日(2024.8.26)		2/F Baidu Campus, N o.10, Shangdi 10th Street, Haidian Dis trict, Beijing 1000
(31)優先権主張番号	202310802054.9		最終頁に続く
(32)優先日	令和5年6月30日(2023.6.30)		
(33)優先権主張国・地域又は機関	中国(CN)		

(54)【発明の名称】 対話モデルのトレーニング方法、回答情報生成方法、装置及び媒体

(57)【特許請求の範囲】

【請求項1】

対話モデルのトレーニング方法であって、

少なくとも一つの第1サンプルデータ及び少なくとも一つの第2サンプルデータを含む第1サンプルデータセットを取得し、前記少なくとも一つの第1サンプルデータにおけるそれぞれは、第1質問テキスト及び第1回答テキストを含み、前記少なくとも一つの第2サンプルデータにおけるそれぞれは、第2質問テキストを含むことと、

前記第1サンプルデータセットを利用して、前記対話モデルをトレーニングするように、

前記少なくとも一つの第1サンプルデータに対応する少なくとも一つの第1質問テキストをそれぞれ前記対話モデルに入力して、前記対話モデルによって出力された対応する少なくとも一つの第1回答予測結果を取得し、

前記少なくとも一つの第2サンプルデータにおけるそれぞれに対する第2質問テキストに対して、

該第2質問テキストを前記対話モデルに入力して、前記対話モデルによって出力された第2回答予測結果を取得し、

前記第2回答予測結果を、少なくとも一つのサンプル質問と、前記少なくとも一つのサンプル質問におけるそれぞれに対応する複数の回答テキストと、前記複数の回答テキストにおけるそれぞれのラベルとに基づいてトレーニングして取得される奨励モデルに入力して前記奨励モデルにより出力された前記第2回答予測結果のスコアを取得し、前記ラベルが、対応する回答テキストのユーザ満足度を示すという操作を実行し、

10

20

前記少なくとも一つの第 1 回答予測結果と、前記少なくとも一つの第 1 サンプルデータにおけるそれぞれの第 1 回答テキストと、前記少なくとも一つの第 2 サンプルデータにおけるそれぞれに対応するスコアとに基づいて、総合損失を確定し、

前記総合損失に基づいて、前記対話モデルの少なくとも一つのパラメータを調整するという第 1 トレーニングプロセスを実行することと、を含む対話モデルのトレーニング方法。

【請求項 2】

前記少なくとも一つの第 1 回答予測結果と、前記少なくとも一つの第 1 サンプルデータにおけるそれぞれの第 1 回答テキストと、前記少なくとも一つの第 2 サンプルデータにおけるそれぞれに対応するスコアとに基づいて、総合損失を確定することは、

前記少なくとも一つの第 1 サンプルデータにおけるそれぞれの第 1 回答テキスト及び対応する第 1 回答予測結果に基づいて、第 1 損失を確定することと、

前記少なくとも一つの第 2 サンプルデータに対応する少なくとも一つのスコアに基づいて、第 2 損失を確定することと、

前記第 1 損失及び前記第 2 損失に基づいて、前記総合損失を確定することと、を含む、請求項 1 に記載の方法。

【請求項 3】

前記少なくとも一つの第 2 サンプルデータに対応する少なくとも一つのスコアに基づいて、第 2 損失を確定することは、

前記少なくとも一つのスコアに基づいて、前記少なくとも一つのスコアの平均及び分散を確定することと、

前記少なくとも一つのスコアにおける各スコアに対して、前記平均及び前記分散に基づいて該スコアを正規化して、更新されたスコアを取得することと、

少なくとも一つの更新されたスコアに基づいて前記第 2 損失を確定することと、を含む、請求項 2 に記載の方法。

【請求項 4】

前記第 1 損失及び前記第 2 損失に基づいて、前記総合損失を確定することは、

前記第 1 損失に対応する第 1 所定重みと、前記第 2 損失に対応する第 2 所定重みに基づいて、前記第 1 損失と前記第 2 損失とを重み付けて、前記総合損失を取得することを含む、請求項 3 に記載の方法。

【請求項 5】

事前トレーニング済み言語モデルと、少なくとも一つの第 3 サンプルデータを含む第 2 サンプルデータセットとを取得し、前記少なくとも一つの第 3 サンプルデータにおけるそれぞれは、第 3 質問テキスト及び第 3 回答テキストを含み、前記事前トレーニング済み言語モデルは所定数量の教師なしサンプルコーパスに基づいてトレーニングして取得されることと、

前記第 1 サンプルデータセットを利用して前記対話モデルをトレーニングする前に、初期対話モデルを取得するように、前記第 2 サンプルデータセットにおける各第 3 サンプルデータに基づいて、前記事前トレーニング済み言語モデルが収束するまで、

該第 3 サンプルデータに対応する第 3 質問テキストを前記事前トレーニング済み言語モデルに入力して、前記事前トレーニング済み言語モデルによって出力された第 3 回答予測結果を取得し、

前記第 3 回答予測結果と該第 3 サンプルデータに対応する第 3 回答テキストとに基づいて、前記事前トレーニング済み言語モデルのパラメータを調整して、前記事前トレーニング済み言語モデルを更新するという前記事前トレーニング済み言語モデルに対するトレーニング操作を繰り返して実行することをさらに含む、請求項 4 に記載の方法。

【請求項 6】

前記対話モデルは、前記初期対話モデルをもとに少なくとも一回の前記第 1 トレーニングプロセスを経て取得されるものであり、前記方法は、

該第 2 質問テキストを前記初期対話モデルに入力して、前記初期対話モデルによって出

10

20

30

40

50

力された第4回答予測結果を取得することをさらに含み、

前記少なくとも一つの第2サンプルデータに対応する少なくとも一つのスコアに基づいて、第2損失を確定することは、

前記少なくとも一つのスコアと、前記少なくとも一つの第2サンプルデータにおけるそれぞれに対応する第2質問テキストと、第2回答予測結果と、第4回答予測結果とに基づいて、前記第2損失を確定することを含む、請求項5に記載の方法。

【請求項7】

前記少なくとも一つの第1サンプルデータの第1数と、前記少なくとも一つの第2サンプルデータの第2数とは、それぞれ複数であり、且つ前記第1数と前記第2数とは、所定割合に符合する、請求項6に記載の方法。

【請求項8】

対話モデルに基づいて実現される回答情報生成方法であって、

ユーザの質問テキストを取得することと、

前記質問テキストを前記対話モデルに入力して、前記対話モデルが生成した回答テキストを取得することを含み、前記対話モデルは、請求項1～7のいずれか一項に記載のトレーニング方法に従ってトレーニングされたものである、対話モデルに基づいて実現される回答情報生成方法。

【請求項9】

対話モデルのトレーニング装置であって、

少なくとも一つの第1サンプルデータ及び少なくとも一つの第2サンプルデータを含む第1サンプルデータセットを取得するように構成される第1取得ユニットであって、前記少なくとも一つの第1サンプルデータにおけるそれぞれは、第1質問テキスト及び第1回答テキストを含み、前記少なくとも一つの第2サンプルデータにおけるそれぞれは、第2質問テキストを含むものと、

前記第1サンプルデータセットを利用して、前記対話モデルをトレーニングするように、

前記少なくとも一つの第1サンプルデータに対応する少なくとも一つの第1質問テキストをそれぞれ前記対話モデルに入力して、前記対話モデルによって出力された対応する少なくとも一つの第1回答予測結果を取得し、

前記少なくとも一つの第2サンプルデータにおけるそれぞれに対する第2質問テキストに対して、

該第2質問テキストを前記対話モデルに入力して、前記対話モデルによって出力された第2回答予測結果を取得し、

前記第2回答予測結果を、少なくとも一つのサンプル質問と、前記少なくとも一つのサンプル質問におけるそれぞれに対応する複数の回答テキストと、前記複数の回答テキストにおけるそれぞれのラベルとに基づいてトレーニングして取得される奨励モデルに入力して前記奨励モデルにより出力された第2回答予測結果のスコアを取得し、前記ラベルが、対応する回答テキストのユーザ満足度を示すという操作を実行し、

前記少なくとも一つの第1回答予測結果と、前記少なくとも一つの第1サンプルデータにおけるそれぞれの第1回答テキストと、前記少なくとも一つの第2サンプルデータにおけるそれぞれに対応するスコアとに基づいて、総合損失を確定し、

前記総合損失に基づいて、前記対話モデルの少なくとも一つのパラメータを調整するという第1トレーニングプロセスを実行するように構成される第1トレーニングユニットとを含む、対話モデルのトレーニング装置。

【請求項10】

前記少なくとも一つの第1回答予測結果と、前記少なくとも一つの第1サンプルデータにおけるそれぞれの第1回答テキストと、前記少なくとも一つの第2サンプルデータにおけるそれぞれに対応するスコアとに基づいて、総合損失を確定することは、

前記少なくとも一つの第1サンプルデータにおけるそれぞれの第1回答テキスト及び対応する第1回答予測結果に基づいて、第1損失を確定することと、

前記少なくとも一つの第2サンプルデータに対応する少なくとも一つのスコアに基づい

10

20

30

40

50

て、第 2 損失を確定することと、

前記第 1 損失及び前記第 2 損失に基づいて、前記総合損失を確定することと、を含む、請求項 9 に記載の装置。

【請求項 1 1】

前記少なくとも一つの第 2 サンプルデータに対応する少なくとも一つのスコアに基づいて、第 2 損失を確定することは、

前記少なくとも一つのスコアに基づいて、前記少なくとも一つのスコアの平均及び分散を確定することと、

前記少なくとも一つのスコアにおける各スコアに対して、前記平均及び前記分散に基づいて該スコアを正規化して、更新されたスコアを取得することと、

少なくとも一つの更新されたスコアに基づいて前記第 2 損失を確定することと、を含む、請求項 1 0 に記載の装置。

【請求項 1 2】

前記第 1 損失及び前記第 2 損失に基づいて、前記総合損失を確定することは、

前記第 1 損失に対応する第 1 所定重みと、前記第 2 損失に対応する第 2 所定重みとに基づいて、前記第 1 損失と前記第 2 損失とを重み付けて、前記総合損失を取得することを含む、請求項 1 1 に記載の装置。

【請求項 1 3】

事前トレーニング済み言語モデルと、少なくとも一つの第 3 サンプルデータを含む第 2 サンプルデータセットとを取得するように構成される第 2 取得ユニットであって、前記少なくとも一つの第 3 サンプルデータにおけるそれぞれは、第 3 質問テキスト及び第 3 回答テキストを含み、前記事前トレーニング済み言語モデルは所定数量の教師なしサンプルコーパスに基づいてトレーニングされたものと、

前記第 1 サンプルデータセットを利用して前記対話モデルをトレーニングする前に、初期対話モデルを取得するように、前記第 2 サンプルデータセットにおける各第 3 サンプルデータに基づいて、前記事前トレーニング済み言語モデルが収束するまで、

該第 3 サンプルデータに対応する第 3 質問テキストを前記事前トレーニング済み言語モデルに入力して、前記事前トレーニング済み言語モデルによって出力された第 3 回答予測結果を取得し、

前記第 3 回答予測結果と該第 3 サンプルデータに対応する第 3 回答テキストとに基づいて、前記事前トレーニング済み言語モデルのパラメータを調整して、前記事前トレーニング済み言語モデルを更新するという前記事前トレーニング済み言語モデルに対するトレーニング操作を繰り返して実行するように構成される第 2 トレーニングユニットとを含む、請求項 1 2 に記載の装置。

【請求項 1 4】

前記対話モデルは、前記初期対話モデルをもとに少なくとも一回の前記第 1 トレーニングプロセスを経て取得されるものであり、前記装置は、

該第 2 質問テキストを前記初期対話モデルに入力して、前記初期対話モデルによって出力された第 4 回答予測結果を取得するように構成される第 3 取得ユニットをさらに含み、

前記少なくとも一つの第 2 サンプルデータに対応する少なくとも一つのスコアに基づいて、第 2 損失を確定することは、

前記少なくとも一つのスコアと、前記少なくとも一つの第 2 サンプルデータにおけるそれぞれに対応する第 2 質問テキストと、第 2 回答予測結果と、第 4 回答予測結果とに基づいて、前記第 2 損失を確定することを含む、請求項 1 3 に記載の装置。

【請求項 1 5】

前記少なくとも一つの第 1 サンプルデータの第 1 数と、前記少なくとも一つの第 2 サンプルデータの第 2 数とは、それぞれ複数であり、且つ前記第 1 数と前記第 2 数とは、所定割合に符合する、請求項 1 4 に記載の装置。

【請求項 1 6】

対話モデルに基づいて実現される回答情報生成装置であって、

10

20

30

40

50

ユーザの質問テキストを取得するように構成される取得ユニットと、
前記質問テキストを前記対話モデルに入力して、前記対話モデルが生成した回答テキストを取得するように構成される生成ユニットと、を含み、前記対話モデルは、請求項 1 ~ 7 のいずれか一項に記載のトレーニング方法に従ってトレーニングされたものである、対話モデルに基づいて実現される回答情報生成装置。

【請求項 17】

電子機器であって、
少なくとも一つのプロセッサと、
前記少なくとも一つのプロセッサに通信接続されたメモリと、を含み、
前記メモリには、前記少なくとも一つのプロセッサによって実行可能な命令が記憶されており、前記命令が前記少なくとも一つのプロセッサにより実行されて、前記少なくとも一つのプロセッサに、請求項 1 ~ 7 のいずれか一項に記載の方法を実行させることを可能にする、電子機器。

10

【請求項 18】

コンピュータ命令を記憶した非一時的コンピュータ可読記憶媒体であって、前記コンピュータ命令は、コンピュータに請求項 1 ~ 7 のいずれか一項に記載の方法を実行させるために用いられる、非一時的コンピュータ可読記憶媒体。

【請求項 19】

プロセッサにより実行されると、請求項 1 ~ 7 のいずれか一項に記載の方法を実行するコンピュータプログラム。

20

【発明の詳細な説明】

【技術分野】

【0001】

本開示は人工知能技術分野に関し、特に自然言語処理、知能対話技術分野に関し、具体的に対話モデルのトレーニング方法、対話モデルに基づいて実現される回答情報生成方法、装置、電子機器、コンピュータ可読記憶媒体及コンピュータプログラムに関する。

【背景技術】

【0002】

人工知能は、コンピュータに人間のいくつかの思惟過程及び知的行動（例えば、学習、推理、思考、計画など）を模擬させるように研究する科目であり、ハードウェア面の技術もあれば、ソフトウェア面の技術もある。人工知能のハードウェア技術は、一般的にセンサ、人工知能専用チップ、クラウドコンピューティング、分散ストレージ、ビッグデータ処理などの技術を含み、人工知能ソフトウェア技術は、主にコンピュータ視覚技術、音声認識技術、自然言語処理技術及び機械学習/深層学習、ビッグデータ処理技術、ナレッジグラフ技術などのいくつかの大きな方向を含む。

30

【0003】

超大規模言語モデルに基づくタスク型対話生成技術は、現在人工知能分野の研究の焦点の一つである。この技術は、大規模言語モデルの自然言語生成能力を利用して、タスク型対話の特定のニーズに結びつけて特定のタスクの要件を満たす対話コンテンツを生成できる。

40

【0004】

該部分で説明される方法は、必ずしも以前に想定された方法又は採用された方法ではない。特に断りのない限り、該部分に記載されているいずれの方法は、該部分に含まれるだけで従来技術であると考えられるべきではない。同様に、特に断りのない限り、該部分で言及されている課題は、従来の技術で承認されたものであると考えるべきではない。

【発明の概要】

【0005】

本開示は、対話モデルのトレーニング方法、対話モデルに基づいて実現される回答情報生成方法、装置、電子機器、コンピュータ可読記憶媒体及コンピュータプログラムを提供する。

50

【0006】

本開示の一態様によれば、対話モデルのトレーニング方法を提供し、少なくとも一つの第1サンプルデータ及び少なくとも一つの第2サンプルデータを含む第1サンプルデータセットを取得し、少なくとも一つの第1サンプルデータにおけるそれぞれは、第1質問テキスト及び第1回答テキストを含み、少なくとも一つの第2サンプルデータにおけるそれぞれは、第2質問テキストを含むことと、第1サンプルデータセットを利用して、対話モデルをトレーニングするように、少なくとも一つの第1サンプルデータに対応する少なくとも一つの第1質問テキストをそれぞれ対話モデルに入力して、対話モデルによって出力された対応する少なくとも一つの第1回答予測結果を取得し、少なくとも一つの第2サンプルデータにおけるそれぞれに対する第2質問テキストに対して、該第2質問テキストを対話モデルに入力して、対話モデルによって出力された第2回答予測結果を取得し、第2回答予測結果を、少なくとも一つのサンプル質問と、少なくとも一つのサンプル質問におけるそれぞれに対応する複数の回答テキストと、複数の回答テキストにおけるそれぞれのラベルとに基づいてトレーニングして取得される奨励モデルに入力して奨励モデルにより出力された第2回答予測結果のスコアを取得し、ラベルが、対応する回答テキストのユーザ満足度を示すという操作を実行し、少なくとも一つの第1回答予測結果と、少なくとも一つの第1サンプルデータにおけるそれぞれの第1回答テキストと、少なくとも一つの第2サンプルデータにおけるそれぞれに対応するスコアとに基づいて、総合損失を確定し、総合損失に基づいて、対話モデルの少なくとも一つのパラメータを調整するという第1トレーニングプロセスを実行することと、を含む。

10

20

【0007】

本開示の一態様によれば、対話モデルに基づいて実現される回答情報生成方法を提供し、ユーザの質問テキストを取得することと、質問テキストを対話モデルに入力して、対話モデルが生成した回答テキストを取得することを含み、対話モデルは、上記対話モデルのトレーニング方法に従ってトレーニングされたものである。

【0008】

本開示の一態様によれば、対話モデルのトレーニング装置を提供し、少なくとも一つの第1サンプルデータ及び少なくとも一つの第2サンプルデータを含む第1サンプルデータセットを取得するように構成される第1取得ユニットであって、少なくとも一つの第1サンプルデータにおけるそれぞれは、第1質問テキスト及び第1回答テキストを含み、少なくとも一つの第2サンプルデータにおけるそれぞれは、第2質問テキストを含むものと、第1サンプルデータセットを利用して、対話モデルをトレーニングするように、少なくとも一つの第1サンプルデータに対応する少なくとも一つの第1質問テキストをそれぞれ対話モデルに入力して、対話モデルによって出力された対応する少なくとも一つの第1回答予測結果を取得し、少なくとも一つの第2サンプルデータにおけるそれぞれに対する第2質問テキストに対して、該第2質問テキストを対話モデルに入力して、対話モデルによって出力された第2回答予測結果を取得し、第2回答予測結果を、少なくとも一つのサンプル質問と、少なくとも一つのサンプル質問におけるそれぞれに対応する複数の回答テキストと、複数の回答テキストにおけるそれぞれのラベルとに基づいてトレーニングして取得される奨励モデルに入力して奨励モデルにより出力された第2回答予測結果のスコアを取得し、ラベルが、対応する回答テキストのユーザ満足度を示すという操作を実行し、少なくとも一つの第1回答予測結果と、少なくとも一つの第1サンプルデータにおけるそれぞれの第1回答テキストと、少なくとも一つの第2サンプルデータにおけるそれぞれに対応するスコアとに基づいて、総合損失を確定し、総合損失に基づいて、対話モデルの少なくとも一つのパラメータを調整するという第1トレーニングプロセスを実行するように構成される第1トレーニングユニットとを含む。

30

40

【0009】

本発明の他の態様によれば、対話モデルに基づいて実現される回答情報生成装置を提供し、ユーザの質問テキストを取得するように構成される取得ユニットと、質問テキストを対話モデルに入力して、対話モデルが生成した回答テキストを取得するように構成される

50

生成ユニットとを含み、ここで、対話モデルは、上記対話モデルのトレーニング方法に従ってトレーニングされたものである。

【0010】

本開示の別の態様によれば、電子機器を提供し、少なくとも一つのプロセッサと、少なくとも一つのプロセッサに通信接続されたメモリとを含み、ここでメモリには、少なくとも一つのプロセッサによって実行可能な命令が記憶されており、命令が少なくとも一つのプロセッサにより実行されて、少なくとも一つのプロセッサに、上記対話モデルのトレーニング方法又は対話モデルに基づいて実現される回答情報生成方法を実行させることを可能にする。

【0011】

本開示の別の態様によれば、コンピュータ命令を記憶した非一時的コンピュータ可読記憶媒体を提供し、ここで、コンピュータ命令は、コンピュータに、上記対話モデルのトレーニング方法又は対話モデルに基づいて実現される回答情報生成方法を実行させるために用いられる。

【0012】

本開示の別の態様によれば、プロセッサにより実行されると、上記対話モデルのトレーニング方法又は対話モデルに基づいて実現される回答情報生成方法を実現するコンピュータプログラムを提供する。

【0013】

本開示の一つ又は複数の実施例によれば、対話モデルの人工フィードバックに基づく強化学習トレーニング段階において、教師あり微調整トレーニングの損失を導入することにより、強化学習段階において、教師あり微調整トレーニング中に学習された対話タスクを解決する能力が忘れられないようにし、対話モデルの事実精度及びユーザ意図の理解能力を向上させ、これにより、対話モデルの回答情報生成効果を全体的に向上させることができる。

【0014】

理解すべきこととして、該部分に説明される内容は、本開示の実施例の要点又は重要な特徴を識別することを意図しておらず、本開示の保護範囲を限定するためのものではない。本開示の他の特徴は、以下の明細書によって容易に理解されるであろう。

【図面の簡単な説明】

【0015】

図面は、実施例を例示的に示し、明細書の一部を構成し、明細書の文字による説明とともに、実施例の例示的な実施形態を説明するために用いられる。図示の実施例は例示的目的のみであり、特許請求の範囲を限定するものではない。全ての図面において、同一の符号は、類似しているが、必ずしも同じとは限らない要素を指す。

【0016】

【図1】本開示の実施例による、本明細書で説明される各方法を実施することができる例示的なシステムの概略図を示す。

【図2】本開示の実施例による、対話モデルのトレーニング方法のフローチャートを示す。

【図3】本開示の実施例による、総合損失を確定するフローチャートを示す。

【図4】本開示の実施例による、第2損失を確定するフローチャートを示す。

【図5】本開示の実施例による、対話モデルのトレーニング方法のフローチャートを示す。

【図6】本開示の実施例による、対話モデルに基づいて実現される回答情報生成方法のフローチャートを示す。

【図7】本開示の実施例による、対話モデルのトレーニング装置の構成ブロック図を示す。

【図8】本開示の実施例による、対話モデルに基づいて実現される回答情報生成装置の構成ブロック図を示す。

【図9】本開示の実施例を実現するための例示的な電子機器の構成ブロック図を示す。

【発明を実施するための形態】

【0017】

10

20

30

40

50

以下、図面に合わせて本開示の例示的な実施例を説明して、それに含まれる本開示の実施例における様々な詳細が理解を助けるため、それらは単なる例示的なものと考えられるべきである。従って、当業者であれば、本開示の範囲及び精神から逸脱することなく、本明細書で説明された実施例に対して様々な変更及び修正を行うことができることを認識すべきである。同様に、明瞭と簡潔のために、以下の説明では公知の機能及び構造についての説明を省略している。

【 0 0 1 8 】

本開示では、特に明記しない限り、様々な要素を説明するための「第 1」、「第 2」などの用語は、これらの要素の位置関係、タイミング関係、又は重要性関係を限定することを意図していない。このような用語は、一要素を別の要素から区別するためにのみ使用される。いくつかの例では、第 1 の要素と第 2 要素は、要素の同じ例を指してもよく、場合によっては、コンテキストの説明に基づいて、異なる例を指してもよい。

【 0 0 1 9 】

本開示の様々な例の説明で使用される用語は、特定の例を説明することのみを目的としており、限定することを意図していない。コンテキストで別途に明確に示されていない限り、特に要素の数を限定しないなら、要素は一つであってもよいし、複数であってもよい。なお、本開示で使用される用語「及び/又は」は、リストされたアイテムのいずれか及び可能な全ての組み合わせをカバーする。

【 0 0 2 0 】

以下、図面を参照して本開示の実施例について詳細に説明する。

【 0 0 2 1 】

図 1 は、本開示の実施例によると、本明細書に記載された様々な方法及び装置を、その中で実施することができる例示的なシステム 100 の概略図を示す。図 1 を参照すると、該システム 100 は、一つ又は複数のクライアントデバイス 101、102、103、104、105 と 106、サーバ 120、及び一つ又は複数のクライアントデバイスをサーバ 120 に結合する一つ又は複数の通信ネットワーク 110 を含む。クライアントデバイス 101、102、103、104、105 と 106 は、一つ又は複数のアプリケーションを実行するように構成されることが可能である。

【 0 0 2 2 】

本開示の実施例では、サーバ 120 は、上記対話モデルのトレーニング方法又は対話モデルに基づいて実現される回答情報生成方法を実行できるようにする一つ又は複数のサービス又はソフトウェアアプリケーションを実行することができる。

【 0 0 2 3 】

いくつかの実施例では、サーバ 120 は、非仮想環境と仮想環境を含むことができる他のサービス又はソフトウェアアプリケーションも提供することができる。いくつかの実施例では、これらのサービスは、web ベースのサービス又はクラウドサービスとして提供することができ、例えば、ソフトウェアアズサービス (SaaS) モデルでクライアントデバイス 101、102、103、104、105 及び/又は 106 のユーザに提供される。

【 0 0 2 4 】

図 1 に示す構成では、サーバ 120 は、サーバ 120 により実行される機能を実現する一つ又は複数のアセンブリを含んでもよい。これらのアセンブリは、一つ又は複数のプロセッサで実行できるソフトウェアアセンブリ、ハードウェアアセンブリ、又はそれらの組み合わせを含んでもよい。クライアントデバイス 101、102、103、104、105 及び/又は 106 を操作するユーザは、これらのアセンブリが提供するサービスを利用するために、一つ又は複数のクライアントアプリケーションを順次利用してサーバ 120 とやり取りをすることができる。様々な異なるシステム構成が可能であり、システム 100 とは異なってもよいことを理解されたい。したがって、図 1 は、本明細書に記載された様々な方法を実施するためのシステムの一例であり、制限することを意図していない。

【 0 0 2 5 】

ユーザは、クライアントデバイス101、102、103、104、105、及び/又は106を使用して、対話テキストを入力することができる。クライアントデバイスは、クライアントデバイスのユーザがクライアントデバイスとやり取りするインターフェースを提供することができる。クライアントデバイスは、該インターフェースを介してユーザに情報を出力することもできる。図1では6つのクライアントデバイスしか図示されていないが、当業者であれば理解できるように、本開示はいかなる数のクライアントデバイスもサポートできる。

【0026】

クライアントデバイス101、102、103、104、105及び/又は106は、携帯型ハンドヘルドデバイス、汎用コンピュータ（例えば、パーソナルコンピュータやノートパソコン）、ワークステーションコンピュータ、ウェアラブルデバイス、スマートスクリーンデバイス、セルフサービス端末デバイス、サービスロボット、ゲームシステム、シンクライアント、各種のメッセージングデバイス、センサ、又はその他の検知デバイスなどの様々なタイプのコンピュータデバイスを含んでもよい。これらのコンピュータデバイスは、MICROSOFT Windows、APPLE iOS、類UNIX（登録商標）オペレーティングシステム、Linux（登録商標）又は類Linux（登録商標）オペレーティングシステム（例えば、GOOGLE Chrome OS）などの様々なタイプ及びバージョンのソフトウェアアプリケーションやオペレーティングシステムを実行したり、MICROSOFT Windows Mobile OS、iOS、Windows Phone、Androidなどの各種のモバイルオペレーティングシステムを含んだりすることができる。携帯用ハンドヘルドデバイスには、携帯電話、インテリジェントフォン、タブレット、パーソナルデジタルアシスタント（PDA）などを含んでもよい。ウェアラブルデバイスは、ヘッドマウント型ディスプレイ（例えば、スマートグラス）と他のデバイスを含んでもよい。ゲームシステムは、様々なハンドヘルド型のゲームデバイス、インターネット対応のゲームデバイスなどを含んでもよい。クライアントデバイスは、例えば、インターネットInternet関連アプリケーション、通信アプリケーション（例えば、電子メールアプリケーション）、ショートメッセージサービス（SMS）アプリケーション、様々なアプリケーションを実行でき、且つ様々な通信プロトコルを使用できる。

【0027】

ネットワーク110は、当業者に知られている任意のタイプのネットワークであってもよく、それは、データ通信をサポートするために、複数の利用可能なプロトコルのいずれか一つ（TCP/IP、SNA、IPXなどを含むがこれらに限定されない）を使用することができる。例として、一つ又は複数のネットワーク110は、ローカルエリアネットワーク（LAN）、イーサネットベースのネットワーク、トークンリング、ワイドエリアネットワーク（WAN）、インターネット、仮想ネットワーク、仮想プライベートネットワーク（VPN）、イントラネット、エクストラネット、ブロックチェーンネットワーク、公衆交換電話網（PSTN）、赤外線ネットワーク、無線ネットワーク（例えば、Bluetooth（登録商標）、WIFI）、及び/又はこれら及び/又はその他のネットワークの任意の組み合わせであってもよい。

【0028】

サーバ120は、一つ又は複数の汎用コンピュータ、専用サーバコンピュータ（例えば、PC（パーソナルコンピュータ）サーバ、UNIX（登録商標）サーバ、ミッドレンジサーバ）、ブレードサーバ、大型コンピュータ、サーバクラスタ、又はその他のいかなる適切な構成及び/又は組み合わせを含んでもよい。サーバ120は、仮想オペレーティングシステムを実行する一つ又は複数の仮想マシン、又は仮想化に関わる他のコンピューティングアーキテクチャ（例えば、サーバの仮想記憶デバイスを維持するために仮想化された論理記憶デバイスの一つ又は複数のフレキシブルプール）を含んでもよい。様々な実施例では、サーバ120は、以下に説明する機能を提供する一つ又は複数のサービス又はソフトウェアアプリケーションを実行することができる。

【 0 0 2 9 】

サーバ120における計算ユニットは、上記した任意のオペレーティングシステム及び任意の商用サーバオペレーティングシステムを含む一つ又は複数のオペレーティングシステムを実行することができる。サーバ120は、HTTPサーバ、FTPサーバ、CGIサーバ、JAV A（登録商標）サーバ、データベースサーバなど、様々な追加のサーバアプリケーション及び/又は中間層アプリケーションのいずれか一つを実行することもできる。

【 0 0 3 0 】

いくつかの実施例では、サーバ120は、クライアントデバイス101、102、103、104、105及び/又は106のユーザから受信したデータフィールド及び/又はイベントの更新を分析及び総合するための一つ又は複数のアプリケーションを含んでもよい。サーバ120は、クライアントデバイス101、102、103、104、105及び/又は106の一つ又は複数のディスプレイデバイスを介してデータフィールド及び/又はリアルタイムイベントを表示する一つ又は複数のアプリケーションを含んでもよい。

10

【 0 0 3 1 】

いくつかの実施例では、サーバ120は、分散型システムのサーバであってもよいし、ブロックチェーンを組み込んだサーバであってもよい。サーバ120は、クラウドサーバであってもよいし、人工知能技術を備えたインテリジェントクラウドコンピューティングサーバやインテリジェントクラウドホストであってもよい。クラウドサーバはクラウドコンピューティングサービスシステムにおけるホスト製品であり、従来の物理ホストと仮想専用サーバ（VPS、Virtual Private Server）サービスに存在する管理難度が大きく、業務拡張性が弱いという欠陥を解決する。

20

【 0 0 3 2 】

システム100は、一つ又は複数のデータベース130を含むこともできる。いくつかの実施例では、これらのデータベースはデータやその他の情報を記憶するために使用できる。例えば、データベース130のうちの一つ又は複数は、音声ファイルや動画ファイルのような情報を記憶するために使用できる。データベース130は、さまざまな位置に構成することができる。例えば、サーバ120が使用するデータベースは、サーバ120のローカルにあってもよいし、サーバ120から離れて、ネットワーク又は専用の接続を介してサーバ120と通信してもよい。データベース130は、さまざまなタイプであってもよい。いくつかの実施例では、サーバ120が使用するデータベースは、リレーショナルデータベースであってもよい。これらのデータベースのうちの一つ又は複数は、命令に応じてデータベースとデータベースからのデータを記憶、更新、検索できる。

30

【 0 0 3 3 】

いくつかの実施例では、データベース130のうちの一つ又は複数は、アプリケーションによって使用され、アプリケーションのデータを記憶することもできる。アプリケーションで使用されるデータベースは、キー値リポジトリ、オブジェクトリポジトリ、ファイルシステムでサポートされる汎用リポジトリなど、様々なタイプのデータベースであってもよい。

【 0 0 3 4 】

図1のシステム100は、本開示に基づいて説明した様々な方法及び装置を応用することができるように、様々な方法で構成し操作することができる。

40

【 0 0 3 5 】

本開示の実施例によれば、図2に示すように、対話モデルのトレーニング方法を提供し、少なくとも一つの第1サンプルデータ及び少なくとも一つの第2サンプルデータを含む第1サンプルデータセットを取得し、少なくとも一つの第1サンプルデータにおけるそれぞれは、第1質問テキスト及び第1回答テキストを含み、少なくとも一つの第2サンプルデータにおけるそれぞれは、第2質問テキストを含むステップS201と、

第1サンプルデータセットを利用して、対話モデルをトレーニングするように、

少なくとも一つの第1サンプルデータに対応する少なくとも一つの第1質問テキスト

50

をそれぞれ対話モデルに入力して、対話モデルによって出力された対応する少なくとも一つの第1回答予測結果を取得するステップS2021と、

少なくとも一つの第2サンプルデータにおけるそれぞれに対する第2質問テキストに対して、

該第2質問テキストを対話モデルに入力して、対話モデルによって出力された第2回答予測結果を取得するステップS2022と、

第2回答予測結果を、少なくとも一つのサンプル質問と、少なくとも一つのサンプル質問におけるそれぞれに対応する複数の回答テキストと、複数の回答テキストにおけるそれぞれのラベルとに基づいてトレーニングして取得される奨励モデルに入力して奨励モデルにより出力された第2回答予測結果のスコアを取得し、ラベルが、対応する回答テキストのユーザ満足度を示すステップS2023と、

10

少なくとも一つの第1回答予測結果と、少なくとも一つの第1サンプルデータにおけるそれぞれの第1回答テキストと、少なくとも一つの第2サンプルデータにおけるそれぞれに対応するスコアとに基づいて、総合損失を確定するステップS2024と、

総合損失に基づいて、対話モデルの少なくとも一つのパラメータを調整するステップS2025との第1トレーニングプロセスを実行するステップS202を含む。

【0036】

これにより、対話モデルの人工フィードバックに基づく強化学習トレーニング段階において、教師あり微調整トレーニングの損失を導入することにより、強化学習段階において、教師あり微調整トレーニング中に学習された対話タスクを解決する能力が忘れられないようにし、対話モデルの事実精度及びユーザの意図理解能力を向上させ、これにより、対話モデルの回答情報生成効果を全体的に向上させることができる。

20

【0037】

いくつかの実施例では、第1サンプルデータセットは、2種のサンプルデータを含むことができ、ここで、各第1サンプルデータは、第1質問テキストと、それに対応する第1回答テキストとを含み、各第2サンプルデータには、第2質問テキストが一つ含まれる。

【0038】

いくつかの実施例では、第2質問テキストは、第1サンプルデータセットにおけるある第1質問テキストと同じであってもよい。いくつかの実施例では、第1サンプルデータセットにおける第1質問テキスト及び第2質問テキストは、互いに異なってもよい。

30

【0039】

いくつかの実施例では、各第1質問テキスト及び第2質問テキストを現在の対話モデルに入力し、且つそれぞれ対応する第1回答予測結果及び第2回答予測結果を取得することができる。そして、各第2回答予測結果とそれに対応する第2質問テキストを予めトレーニングされた奨励モデルに入力して、該奨励モデルにより出力された第2回答予測結果のスコアを取得する。ここでは、第2回答予測結果のスコアは、該回答情報に対するユーザの満足度を示すために使用され得る。

【0040】

いくつかの実施例では、上記奨励モデルは、以下の方法でトレーニングすることによって得ることができる。まず一又は複数のサンプル質問を取得し、且つ各サンプル質問を順に現在の対話モデルに入力し、それにより各サンプル質問に対して複数の回答テキストを生成し、且つ人工マーキングに基づいて各回答テキストのラベルを取得し、ここでは、ラベルは、対応する回答テキストのユーザ満足度を示すことができる。その後、各回答テキスト及びそれに対応するサンプル質問を初期奨励モデルに入力し、それにより該モデルの予測結果を取得する。続いて、該奨励モデルが収束するまで、予測結果及び対応するラベルに基づいて損失を計算し、且つモデルパラメータを調整することができる。

40

【0041】

いくつかの実施例では、上記奨励モデルは、多層パーセプトロン、ニューラルネットワークなどのアーキテクチャに基づいて確立されてもよい。

【0042】

50

いくつかの実施例では、少なくとも一つの第1回答予測結果と、少なくとも一つの第1サンプルデータにおけるそれぞれの第1回答テキストと、少なくとも一つの第2サンプルデータにおけるそれぞれに対応するスコアとに基づいて、総合損失の計算を行うことができる。例えば、第1質問テキストと第2質問テキストとが同じである場合には、各第1回答予測結果とそれに対応する第1回答テキストとの差異を計算した後、対応するスコアに基づいて、該差異の重み係数を確定し、且つ1つ又は複数が重み係数を乗算した差異に基づいて総合損失を計算することができる。

【0043】

いくつかの実施例では、図3に示すように、少なくとも一つの第1回答予測結果と、少なくとも一つの第1サンプルデータにおけるそれぞれの第1回答テキストと、少なくとも一つの第2サンプルデータにおけるそれぞれに対応するスコアとに基づいて、総合損失を確定することは、

10

少なくとも一つの第1サンプルデータにおけるそれぞれの第1回答テキスト及び対応する第1回答予測結果に基づいて、第1損失を確定するステップS301と、

少なくとも一つの第2サンプルデータに対応する少なくとも一つのスコアに基づいて、第2損失を確定するステップS302と、

第1損失及び第2損失に基づいて、総合損失を確定するステップS303とを含むことができる。

【0044】

これにより、2つの部分の損失をそれぞれ計算し、総合損失を得ることで、対話タスク指向の教師あり微調整トレーニング段階と人工フィードバックによる強化学習段階の2段階のモデルの連合モデリングを実現し、相互に促進し、ユーザの好みに対するモデルのモデリング能力を向上させるとともに、ユーザの指示に対するモデルの理解と満足能力を維持し、さらに、対話モデルの回答情報生成効果を全体的に高めることができる。

20

【0045】

いくつかの実施例では、第1損失は、少なくとも一つの第1サンプルデータにおけるそれぞれに対する第1回答予測結果と、それに対応する第1回答テキストとの間の差異に基づいて確定することができる。

【0046】

いくつかの実施例では、交差エントロピー誤差関数又は平均二乗誤差関数に基づいて上記差異をはかることができる。

30

【0047】

いくつかの実施例では、第2損失は、少なくとも一つの第2回答予測結果に対応する少なくとも一つのスコアの平均又は期待値であってもよい。

【0048】

いくつかの実施例では、混合損失、交互最小化損失などの組み合わせによって総合損失を得ることができる。

【0049】

いくつかの実施例では、第1損失及び第2損失に基づいて、総合損失を確定することは、第1損失に対応する第1所定重みと、第2損失に対応する第2所定重みとに基づいて、第1損失及び第2損失に重み付けを行い、総合損失を得ることを含むことができる。

40

【0050】

これにより、2つの損失に対して重み付けを行うことによって、重みによって2つのトレーニング方法によるモデル全体のトレーニングへの影響を制御し、トレーニングの効果を保証する。

【0051】

いくつかの実施例では、対話モデルのトレーニング段階において、複数の第1サンプルデータセットを順次取得し、複数の第1サンプルデータセットに基づいて複数ラウンドのモデルトレーニングを行うことができる。

【0052】

50

いくつかの実施例では、トレーニングラウンド数が増加するにつれて、合計が一定であることを保証する状態で、第1所定重み及び第2所定重みを対应的に調整して、異なるトレーニングラウンドで損失関数の割合を調整することで、モデルトレーニングのプロセスを制御することができる。例えば、変動因子を一つ設定し、且つトレーニングラウンド数の増加に応じて変動因子によって第2所定重みを漸減しながら、第1所定重みを徐々に増加させ、それにより、モデルのトレーニング効果を保証しつつ、モデルのトレーニング効率を向上させることができる。

【0053】

いくつかの実施例では、図4に示すように、少なくとも一つの第2サンプルデータに対応する少なくとも一つのスコアに基づいて、第2損失を確定することは、
 少なくとも一つのスコアに基づいて、少なくとも一つのスコアの平均及び分散を確定するステップS401と、
 少なくとも一つのスコアの各スコアに対して、平均及び分散に基づいて該スコアを正規化して、更新されたスコアを取得するステップS402と、
 少なくとも一つの更新されたスコアに基づいて、第2損失を確定するステップS403とを含む。

10

【0054】

これにより、スコアを正規化してから損失の計算を行うことにより、トレーニングプロセスにおけるスコア分布をさらに最適化することができ、これにより、強化学習段階で他の損失を導入することによる強化学習の損失がモデルに与える影響が弱まるという問題を避けることができ、強化学習の安定性を向上させることができ、連合最適化の枠組みの下で強化学習損失がモデルトレーニングに及ぼす作用を確保する。

20

【0055】

強化学習の段階では、余分な損失を導入して連合トレーニングを行うことで、強化学習の部分での損失の上昇を抑え、さらに、強化学習が本来果たすべき役割を弱める。

【0056】

いくつかの実施例では、第2損失を計算する前に、現在のラウンドの第1サンプルデータセットにおけるすべての第2回答予測結果に対応するスコアに平均と分散の計算を行うことができる。そして、平均及び分散に基づいて今回のラウンドにおける全てのスコアを正規化処理して正規化スコアを得て、正規化された各スコアに基づいて第2損失の計算を行うことができる。

30

【0057】

いくつかの例示的な実施例では、スコア r_i についての正規化操作は、以下の式によって表すことができる。

【数1】

$$r'_i = (r_i - \bar{r}) / std$$

ここで、

【数2】

$$r'_i$$

40

は正規化後のスコアを表し、

【数3】

$$\bar{r}$$

は平均を表し、stdは分散を表す。

【0058】

これにより、該ラウンド内のスコアを一つの動的標準正規分布下にすることができ、こ

50

れにより、強化学習の安定性を高め、連合最適化の枠組みの下で、強化学習の損失は正常に上昇することを確保する。

【0059】

いくつかの実施例では、第2損失は、正規化されたスコアの平均又は期待値であってもよい。

【0060】

いくつかの実施例では、上記対話モデルのトレーニング方法は、事前トレーニング済み言語モデルと、少なくとも一つの第3サンプルデータを含む第2サンプルデータセットとを取得し、少なくとも一つの第3サンプルデータにおけるそれぞれは、第3質問テキスト及び第3回答テキストを含み、事前トレーニング済み言語モデルは所定数量の教師なしサンプルコーパスに基づいてトレーニングして取得されることと、第1サンプルデータセットを利用して対話モデルをトレーニングする前に、初期対話モデルを取得するように、第2サンプルデータセットにおける各第3サンプルデータに基づいて、事前トレーニング済み言語モデルが収束するまで、該第3サンプルデータに対応する第3質問テキストを事前トレーニング済み言語モデルに入力して、事前トレーニング済み言語モデルによって出力された第3回答予測結果を取得し、第3回答予測結果と該第3サンプルデータに対応する第3回答テキストとに基づいて、事前トレーニング済み言語モデルのパラメータを調整して、事前トレーニング済み言語モデルを更新するという事前トレーニング済み言語モデルに対するトレーニング操作を繰り返して実行することとをさらに含むことができる。

【0061】

これにより、第1トレーニングプロセスを行う前に、まず事前トレーニング済み言語モデルに基づいて教師あり微調整トレーニングを行って、初期対話モデルを取得する。その後さらに初期対話モデルに基づいて教師あり微調整トレーニング及び人工フィードバックに基づく強化学習トレーニングの連合トレーニングを行い、それによりモデルに対話タスクを解決する能力を持たせた上で、ユーザの好みに対する予測能力をさらに取得し、これにより、モデルの全体的な性能を向上させる。

【0062】

図5は、本開示の実施例による対話モデルのトレーニング方法のフローチャートを示す。

【0063】

いくつかの実施例によれば、図5に示すように、対話モデルのトレーニングプロセスは、所定規模の教師なしコーパスに基づいて、汎用事前トレーニング済み言語モデルのトレーニングを行って、汎用事前トレーニング済み言語モデルを取得するステップS501と、汎用事前トレーニング済み言語モデルに対して教師あり微調整トレーニングを行い、具体的には、まず、1つの第2サンプルデータセットを取得し、その中の第3サンプルデータを適用し、該汎用事前トレーニング済み言語モデルに対して教師あり微調整トレーニングを行うことにより、ユーザが入力した質問又は指令に含まれる意図を理解でき、且つこの意図に基づいて比較的質の高い回答能力を与える初期対話モデルを得るステップS502と、強化学習に基づいて初期対話モデルをトレーニングして、最終的な対話モデルを得て、ここで、該ステップでは、上述したトレーニング方法に基づいて、複数ロットの第1サンプルデータセットを取得し、各ラウンドにおいて、第1サンプルデータセットにおける2つの異なるサンプルデータを適用することにより、モデルを連合トレーニングして、最終的な対話モデルを得るステップS503とを含むことができる。

【0064】

いくつかの実施例では、第2サンプルデータセットにおける第3サンプルデータ及び第1サンプルデータセットにおける第1サンプルデータは、同一の予め準備されたサンプルデータセットから得ることができ、該サンプルデータセットにおける各サンプルデータは、いずれもサンプル質問及びそれに対応する回答情報を有する。これにより、モデルにユーザの好みをより良く予測させながら、ユーザが入力した質問又は指令に対するその理解能力、及び高品質の回答情報を生成する能力を維持することができる。

【0065】

10

20

30

40

50

いくつかの実施例では、対話モデルは、初期対話モデルをもとに少なくとも一回の第1トレーニングプロセスを経て得たものであり、上記対話モデルのトレーニング方法は、該第2質問テキストを初期対話モデルに入力して、初期対話モデルによって出力された第4回答予測結果を取得することをさらに含み、ここで、少なくとも一つの第2サンプルデータに対応する少なくとも一つのスコアに基づいて、第2損失を確定することは、少なくとも一つのスコアと、少なくとも一つの第2サンプルデータにおけるそれぞれに対応する第2質問テキストと、第2回答予測結果と、第4回答予測結果とに基づいて、第2損失を確定することを含む。

【0066】

これにより、第2損失計算プロセスでは、第2質問テキスト、第2回答予測結果及び第4回答予測結果に基づいて計算された正規化項を導入することにより、モデルトレーニングの安定性をさらに向上させることができる。

10

【0067】

いくつかの実施例では、各第2サンプルデータの第2質問テキストが第2回答予測結果と第4回答予測結果とに対応する差異に基づいて、正規化項を確定することができる。そして、該正規化項及び該第2回答予測結果に対応するスコア（又は正規化されたスコア）に基づいて第2損失を確定する。

【0068】

いくつかの実施例では、上述の正規化項は、KLダイバージェンスに基づいて計算することができる。

20

【0069】

いくつかの実施例では、総合損失は、以下の式によって表すことができる。

【数4】

$$\begin{aligned} \text{objective}(\phi) = & E_{(x, y) \sim D_{\pi_{\phi}^{RL}}} \left[r_{\theta}(x, y) - \beta \log \left(\frac{\pi_{\phi}^{RL}(y|x)}{\pi^{SFT}(y|x)} \right) \right] \\ & + \gamma E_{(s, t) \sim D_{sft}} \left[\log \left(\pi_{\phi}^{RL}(t/s) \right) \right] \end{aligned}$$

ここで、 θ は政策ポリシー関数のパラメータ（すなわち現在の対話モデルのパラメータ）であり、

30

【数5】

$$E_{(x, y) \sim D_{\pi_{\phi}^{RL}}} \left[r_{\theta}(x, y) - \beta \log \left(\frac{\pi_{\phi}^{RL}(y|x)}{\pi^{SFT}(y|x)} \right) \right]$$

は第2損失であり、

【数6】

$$r_{\theta}(x, y)$$

40

は第2質問テキスト x 及び第2回答予測結果 y のスコアであり、

【数7】

$$\log \left(\frac{\pi_{\phi}^{RL}(y|x)}{\pi^{SFT}(y|x)} \right)$$

は正規化項であり、

【数8】

50

$$\pi^{SFT}(y|x)$$

は第2質問テキスト x を初期対話モデルに入力して第2回答予測結果 y を得る確率であり、
【数9】

$$\pi_{\phi}^{RL}(y|x)$$

は現在の対話モデル（少なくとも一つの第1トレーニングプロセスを経て更新された対話モデル）に基づいて第2質問テキスト x を初期対話モデルに入力して第2回答予測結果 y を得る確率であり、 β は、ポリシー更新の規模を制御するためのスーパーパラメータである。 α は重み係数であり、

【数10】

$$E_{(s, t) \sim D_{sft}} [\log (\pi_{\phi}^{RL}(t/s))]]$$

は第1損失であり、

【数11】

$$\log (\pi_{\phi}^{RL}(t/s))$$

は入力 s が与えられた状況で、正解 t を生成する交差エントロピーを表す。

【0070】

いくつかの実施例では、少なくとも一つの第1サンプルデータの第1数と、少なくとも一つの第2サンプルデータの第2数とが、いずれも複数であり、また、第1数及び第2数は、所定割合に符合する。

【0071】

これにより、2種類のサンプルデータの割合を制御することによって、トレーニングプロセスにおいて、連合トレーニングにおける2つのトレーニング方法がモデルに影響する程度を制御することができ、モデルのトレーニング効果を全体的に最適化する。

【0072】

いくつかの実施例では、第1数と第2数との所定割合は、例えば、1:7であってもよい。

【0073】

いくつかの実施例では、複数ラウンドにおける複数の第1サンプルデータセットについて、トレーニングラウンド数の増加につれて第1数と第2数の割合を増加することができ、例えば、第1数の占める割合を徐々に増加する。これにより、モデルのトレーニング効果を保証すると同時に、モデルのトレーニング効率を高めることができる。

【0074】

いくつかの例示的な実施例において、対話モデルは、例えば、対話のための知識拡張大言語モデル（例えばERNIEbot等）に基づいて確立することができる。

【0075】

いくつかの実施例では、図6に示すように、対話モデルに基づいて実現される回答情報生成方法が提供され、ユーザの質問テキストを取得するステップS601と、質問テキストを対話モデルに入力して、対話モデルが生成した回答テキストを取得するステップS602とを含み、ここで、対話モデルは、上記対話モデルのトレーニング方法に従ってトレーニングされたものである。

【0076】

これにより、上記のトレーニング方法でトレーニングして得られた対話モデルを用いて、より優れた事実理解の精度及びユーザ意図の理解能力を備えることができ、それにより

10

20

30

40

50

、ユーザの期待により合致する回答情報を生成することができる。

【0077】

いくつかの実施例では、図7に示すように、対話モデルのトレーニング装置700が提供され、

少なくとも一つの第1サンプルデータ及び少なくとも一つの第2サンプルデータを含む第1サンプルデータセットを取得するように構成される第1取得ユニット710であって、少なくとも一つの第1サンプルデータにおけるそれぞれは、第1質問テキスト及び第1回答テキストを含み、少なくとも一つの第2サンプルデータにおけるそれぞれは、第2質問テキストを含むものと、

第1サンプルデータセットを利用して、対話モデルをトレーニングするように、

少なくとも一つの第1サンプルデータに対応する少なくとも一つの第1質問テキストをそれぞれ対話モデルに入力して、対話モデルによって出力された対応する少なくとも一つの第1回答予測結果を取得し、

少なくとも一つの第2サンプルデータにおけるそれぞれに対する第2質問テキストに対して、

該第2質問テキストを対話モデルに入力して、対話モデルによって出力された第2回答予測結果を取得し、

第2回答予測結果を、少なくとも一つのサンプル質問と、少なくとも一つのサンプル質問におけるそれぞれに対応する複数の回答テキストと、複数の回答テキストにおけるそれぞれのラベルとに基づいてトレーニングして取得される奨励モデルに入力して奨励モデルにより出力された第2回答予測結果のスコアを取得し、ラベルが、対応する回答テキストのユーザ満足度を示すという操作を実行し、

少なくとも一つの第1回答予測結果と、少なくとも一つの第1サンプルデータにおけるそれぞれの第1回答テキストと、少なくとも一つの第2サンプルデータにおけるそれぞれに対応するスコアとに基づいて、総合損失を確定し、

総合損失に基づいて、対話モデルの少なくとも一つのパラメータを調整するという第1トレーニングプロセスを実行するように構成される第1トレーニングユニット720とを含む。

【0078】

いくつかの実施例では、少なくとも一つの第1回答予測結果と、少なくとも一つの第1サンプルデータにおけるそれぞれの第1回答テキストと、少なくとも一つの第2サンプルデータにおけるそれぞれに対応するスコアとに基づいて、総合損失を確定することは、少なくとも一つの第1サンプルデータにおけるそれぞれの第1回答テキスト及び対応する第1回答予測結果に基づいて、第1損失を確定することと、少なくとも一つの第2サンプルデータに対応する少なくとも一つのスコアに基づいて、第2損失を確定することと、第1損失及び第2損失に基づいて、総合損失を確定することと、を含むことができる。

【0079】

いくつかの実施例では、少なくとも一つの第2サンプルデータに対応する少なくとも一つのスコアに基づいて、第2損失を確定することは、少なくとも一つのスコアに基づいて、少なくとも一つのスコアの平均及び分散を確定することと、少なくとも一つのスコアにおける各スコアに対して、平均及び分散に基づいて該スコアを正規化して、更新されたスコアを取得することと、少なくとも一つの更新されたスコアに基づいて第2損失を確定することと、を含むことができる。

【0080】

いくつかの実施例では、第1損失及び第2損失に基づいて、総合損失を確定することは、第1損失に対応する第1所定重みと、第2損失に対応する第2所定重みとに基づいて、第1損失と第2損失とを重み付けて、総合損失を取得することを含むことができる。

【0081】

いくつかの実施例では、上記トレーニング装置は、事前トレーニング済み言語モデルと、少なくとも一つの第3サンプルデータを含む第2サンプルデータセットとを取得するよ

10

20

30

40

50

うに構成される第2取得ユニットであって、少なくとも一つの第3サンプルデータにおけるそれぞれは、第3質問テキスト及び第3回答テキストを含み、事前トレーニング済み言語モデルは所定数の教師なしサンプルコーパスに基づいてトレーニングされたものと、第1サンプルデータセットを利用して対話モデルをトレーニングする前に、初期対話モデルを取得するように、第2サンプルデータセットにおける各第3サンプルデータに基づいて、事前トレーニング済み言語モデルが収束するまで、該第3サンプルデータに対応する第3質問テキストを事前トレーニング済み言語モデルに入力して、事前トレーニング済み言語モデルによって出力された第3回答予測結果を取得し、第3回答予測結果と該第3サンプルデータに対応する第3回答テキストとに基づいて、事前トレーニング済み言語モデルのパラメータを調整して、事前トレーニング済み言語モデルを更新するという事前トレーニング済み言語モデルに対するトレーニング操作を繰り返して実行するように構成される第2トレーニングユニットとをさらに含むことができる。

10

【0082】

いくつかの実施例では、対話モデルは、初期対話モデルに基づいて少なくとも一回の第1トレーニングプロセスを経て取得されたものであり、上記トレーニング装置は、該第2質問テキストを初期対話モデルに入力して、初期対話モデルによって出力された第4回答予測結果を取得するように構成される第3取得ユニットをさらに含み、且つここで、少なくとも一つの第2サンプルデータに対応する少なくとも一つのスコアに基づいて、第2損失を確定することは、少なくとも一つのスコアと、少なくとも一つの第2サンプルデータにおけるそれぞれに対応する第2質問テキストと、第2回答予測結果と、第4回答予測結果とに基づいて、第2損失を確定することを含む。

20

【0083】

いくつかの実施例では、少なくとも一つの第1サンプルデータの第1数と、少なくとも一つの第2サンプルデータの第2数とは、それぞれ複数であり、且つ第1数と第2数とは、所定割合に符合する。

【0084】

いくつかの実施例では、図8に示すように、対話モデルに基づいて実現される回答情報生成装置800がさらに提供され、ユーザの質問テキストを取得するように構成される取得ユニット810と、質問テキストを対話モデルに入力して、対話モデルが生成した回答テキストを取得するように構成される生成ユニット820とを含み、ここで、対話モデルは、上記対話モデルのトレーニング方法に従ってトレーニングされたものである。

30

【0085】

本開示の実施例によれば、電子機器、可読記憶媒体及コンピュータプログラムをさらに提供する。

【0086】

図9を参照して、本開示のサーバ又はクライアントとして機能する電子機器900の構成ブロック図について説明し、これは、本開示の各態様に適用可能なハードウェア装置の一例である。電子機器は、様々な形態のデジタル電子コンピュータ、例えば、ラップトップ型コンピュータ、デスクトップ型コンピュータ、ステージ、パーソナルデジタルアシスタント、サーバ、ブレードサーバ、大型コンピュータ、その他の適切なコンピュータを示す。電子機器は更に、様々な形態の移動装置、例えば、パーソナルデジタル処理、携帯電話、インテリジェントフォン、ウェアラブル機器とその他の類似する計算装置を示してよい。本明細書に示される部品、これらの接続関係及びこれらの機能は例示的なものに過ぎず、本明細書に説明した及び/又は請求した本開示の実現を制限しない。

40

【0087】

図9に示すように、電子機器900は、読取り専用メモリ(ROM)902に記憶されたコンピュータプログラム、又は記憶ユニット908からランダムアクセスメモリ(RAM)903にロードされたコンピュータプログラムに従って、様々な適切な動作及び処理を実行することができる計算ユニット901を含む。また、RAM903には、電子機器900の動作に必要な各種プログラムやデータが記憶されていてもよい。計算ユニット9

50

01、ROM902、RAM903は、バス904を介して互いに接続されている。入力/出力(I/O)インターフェース905もバス904に接続されている。

【0088】

電子機器900における、入力ユニット906、出力ユニット907、記憶ユニット908、及び通信ユニット909を含む複数のコンポーネントは、I/Oインターフェース905に接続される。入力ユニット906は、電子機器900に情報を入力可能な任意のタイプのデバイスであってもよく、入力ユニット906は、入力された数字又は文字情報を受信し、及び電子機器のユーザ設定及び/又は機能制御に関するキー信号入力を生成してもよく、且つマウス、キーボード、タッチスクリーン、トラックボード、トラックボール、操作レバー、マイク、及び/又はリモコンを含むがこれらに限定されない。出力ユニット907は、情報を提示することが可能な任意のタイプの機器であってもよく、且つディスプレイ、スピーカ、動画/音声出力端末、パイプレータ、及び/又はプリンタを含んでもよいが、これらに限定されない。記憶ユニット908は、磁気ディスク、光ディスクを含んでもよいが、これらに限定されない。通信ユニット909は、電子機器900がインターネット等のコンピュータネットワーク及び/又は各種の電気通信ネットワークを介して他の機器と情報/データを交換することを可能にし、また、モデム、ネットワークカード、赤外線通信装置、無線通信トランシーバ、及び/又はチップセット(例えば、Bluetooth装置、802.11装置、WiFi装置、WiMax装置、セルラ通信装置、及び/又は同様のもの)を含むことができるが、これらに限定されない。

【0089】

計算ユニット901は、処理及び計算能力を有する様々な汎用及び/又は専用処理コンポーネントであってもよい。計算ユニット901のいくつかの例は、中央処理ユニット(CPU)、グラフィックス処理ユニット(GPU)、様々な専用人工知能(AI)計算チップ、機械学習モデルアルゴリズムを実行する様々な計算ユニット、デジタル信号プロセッサ(DSP)、及び任意の適切なプロセッサ、コントローラ、マイクロコントローラなどを含むが、これらに限定されない。計算ユニット901は、上記対話モデルのトレーニング方法又は対話モデルに基づいて実現される回答情報生成方法などの上述した様々な方法及び処理を実行する。例えば、いくつかの実施例では、上記対話モデルのトレーニング方法又は対話モデルに基づいて実現される回答情報生成方法は、記憶ユニット908などの機械可読媒体内に有形的に具現化されるコンピュータソフトウェアプログラムとして実現することができる。いくつかの実施例において、コンピュータプログラムの一部の又は全てはROM902及び/又は通信ユニット909を経由して電子機器900にロード及び/又はインストールされてよい。コンピュータプログラムがRAM903にロードされ、且つ計算ユニット901によって実行されると、上記対話モデルのトレーニング方法又は対話モデルに基づいて実現される回答情報生成方法を実行することができる。代替的に、他の実施例では、計算ユニット901は、他の任意の適切な方法で(例えば、ファームウェアによって)上記対話モデルのトレーニング方法又は対話モデルに基づいて実現される回答情報生成方法を実行するように構成されてもよい。

【0090】

本明細書で上述したシステム及び技術の様々な実施形態は、デジタル電子回路システム、集積回路システム、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)、特定用途向け標準製品(ASSP)、システムオンチップ(SOC)、複雑なプログラマブル論理デバイス(CPLD)、コンピュータハードウェア、ファームウェア、ソフトウェア、及び/又はこれらの組み合わせにおいて実装することができる。これらの様々な実施形態は、一つ又は複数のコンピュータプログラムに実施され、該一つ又は複数のコンピュータプログラムは少なくとも一つのプログラマブルプロセッサを含むプログラマブルシステムで実行し及び/又は解釈してもよく、該プログラマブルプロセッサは専用又は汎用プログラマブルプロセッサであってもよく、記憶システム、少なくとも一つの入力装置、少なくとも一つの出力装置からデータと命令を受信し、データと命令を該記憶システム、該少なくとも一つの入力装置、該少なくとも一つの出力

10

20

30

40

50

装置に送信してよいこと、を含んでもよい。

【 0 0 9 1 】

本開示の方法を実施するプログラムコードは一つ又は複数のプログラミング言語のいかなる組み合わせで書かれてよい。これらのプログラムコードを汎用コンピュータ、特殊目的の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又はコントローラに提供してよく、よってプログラムコードはプロセッサ又はコントローラにより実行される時に流れ図及び/又はブロック図に規定の機能/操作を実施する。プログラムコードは完全に機械で実行してよく、部分的に機械で実行してよく、独立ソフトウェアパッケージとして部分的に機械で実行し且つ部分的に遠隔機械で実行してよく、又は完全に遠隔機械又はサーバで実行してよい。

10

【 0 0 9 2 】

本開示のコンテキストにおいて、機械可読媒体は有形の媒体であってもよく、命令実行システム、装置又はデバイスに使用される又は命令実行システム、装置又はデバイスに結合されて使用されるプログラムを具備又は記憶してよい。機械可読媒体は機械可読信号媒体又は機械可読記憶媒体であってもよい。機械可読媒体は、電子、磁気、光学、電磁、赤外線、又は半導体システム、装置又はデバイス、又は上記内容のいかなる適切な組み合わせを含んでもよいが、これらに限定されない。機械可読記憶媒体のより具体的な例は、一つ又は複数のリード線による電気接続、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 (R A M)、リードオンリーメモリ (R O M)、消去可能なプログラマブル読み出し専用メモリ (E P R O M 又はフラッシュメモリ)、光ファイバー、ポータブルコンパクトディスク読み出し専用メモリ (C D - R O M)、光記憶機器、磁気記憶機器、又は上記内容のいかなる適切な組み合わせを含む。

20

【 0 0 9 3 】

ユーザとのインタラクションを提供するために、コンピュータにはここで説明したシステムと技術を実施してよく、該コンピュータは、ユーザに情報を表示するための表示装置 (例えば、 C R T (陰極線管) 又は L C D (液晶ディスプレイ) 監視モニター)、及びキーボードとポインティング装置 (例えば、マウスやトラックボール) を備え、ユーザは該キーボードと該ポインティング装置を介してコンピュータに入力してよい。その他のタイプの装置は更に、ユーザとのインタラクションを提供するためののものであってもよく、例えば、ユーザに提供するフィードバックはいかなる形態の感覚フィードバック (例えば、視覚フィードバック、聴覚フィードバック、又は触覚フィードバック) であってもよく、いかなる形態 (音入力、音声入力、又は触覚入力を含む) でユーザからの入力を受信してよい。

30

【 0 0 9 4 】

ここで述べたシステムや技術は、バックステージ部材を含む計算システム (例えば、データサーバとして) や、ミドルウェア部材を含む計算システム (例えば、アプリケーションサーバ) や、フロントエンド部材を含む計算システム (例えば、グラフィカルユーザインターフェースやウェブブラウザを有するユーザコンピュータであり、ユーザが、そのグラフィカルユーザインターフェースやウェブブラウザを通じて、それらのシステムや技術の実施形態とのインタラクションを実現できる)、あるいは、それらのバックステージ部材、ミドルウェア部材、あるいはフロントエンド部材の任意の組み合わせからなる計算システムには実施されてもよい。システムの部材は、任意の形式や媒体のデジタルデータ通信 (例えば、通信ネットワーク) により相互に接続されてもよい。通信ネットワークの一例は、ローカルネットワーク (L A N)、広域ネットワーク (W A N) とインターネットを含む。

40

【 0 0 9 5 】

コンピュータシステムは、クライアント側とサーバを含んでもよい。クライアント側とサーバは、一般的に相互に遠く離れ、通常、通信ネットワークを介してインタラクションを行う。互にクライアント側 - サーバという関係を有するコンピュータプログラムを対応するコンピュータで運転することによってクライアント側とサーバとの関係を生成する。

50

サーバは、クラウドサーバであってもよく、分散型システムのサーバでも、又はブロックチェーンと組み合わされたサーバであってもよい。

【0096】

理解すべきこととして、前述した様々な形態のフローを用いて、ステップを改めて順位付け、増加又は削除してよい。例えば、本開示に記載された各ステップは、並列的に実行してもよいし、順次実行してもよいし、異なる順序で実行してもよく、本開示に開示された技術案が所望する結果を実現できれば、本文はこれに限定されないことである。

【0097】

本開示の実施例又は例は図面を参照して説明されたが、上記の方法、システム、及び装置は単なる例示的な実施例又は例であり、本発明の範囲はこれらの実施例又は例によって制限されるものではなく、授權後の特許請求の範囲及びその均等範囲のみによって限定されることを理解されたい。実施例又は例の様々な要素は省略されてもよく、又はそれらの均等要素によって代替されてもよい。なお、各ステップは、本開示で説明した順序とは異なる順序で実行されてもよい。更に、実施例又は例の様々な要素は、様々な方法で組み合わせられてもよい。重要なのは、技術の進化に伴い、ここで説明される多くの要素は、本開示の後に現れる同等の要素に置き換えることができるということである。

10

20

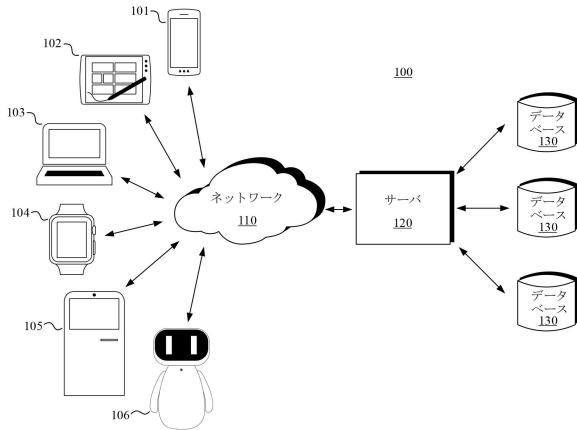
30

40

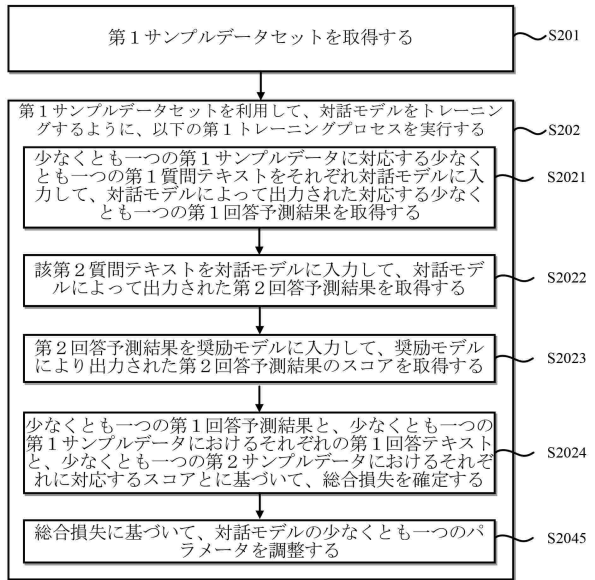
50

【図面】

【図 1】



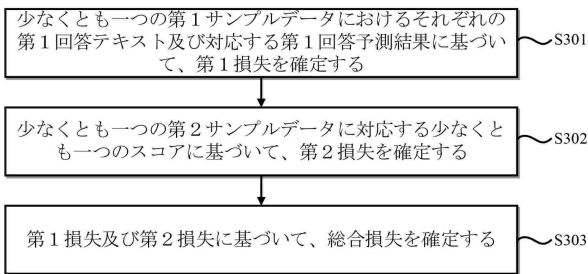
【図 2】



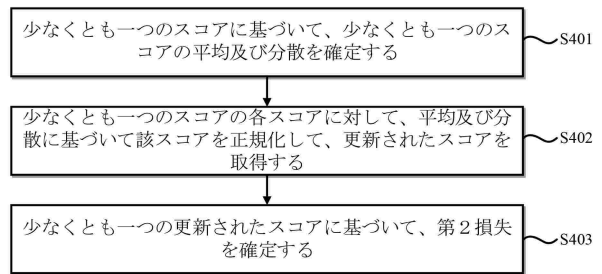
10

20

【図 3】



【図 4】

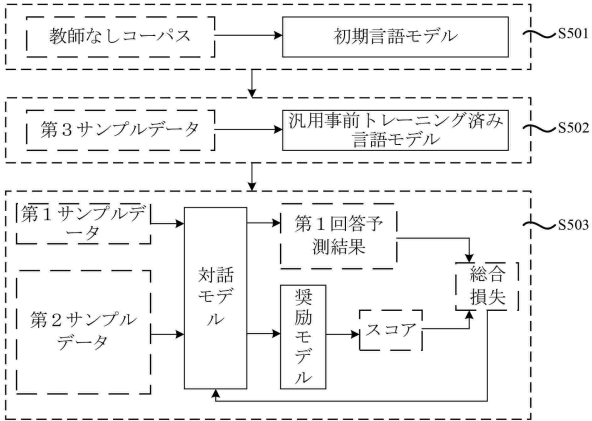


30

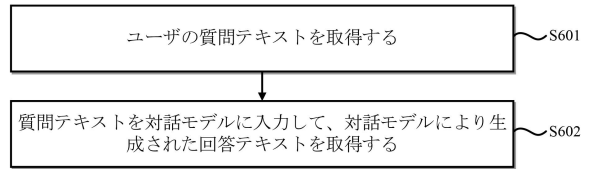
40

50

【図5】

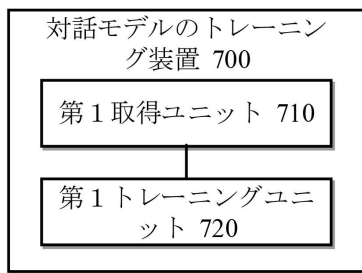


【図6】



10

【図7】

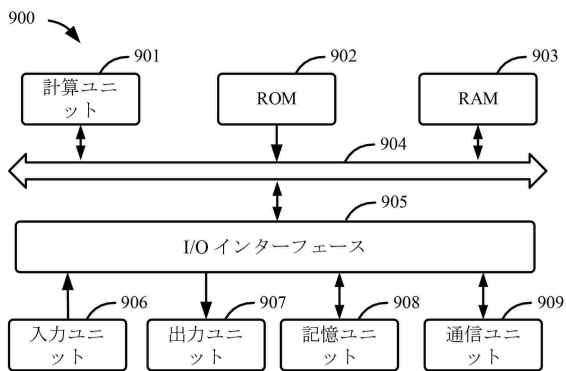


【図8】



20

【図9】



30

40

50

フロントページの続き

- 85, China
- (74)代理人 100107766
弁理士 伊東 忠重
- (74)代理人 100229448
弁理士 中楨 利明
- (72)発明者 イエンピン ジャオ
中華人民共和国, 100085 ベイジン, ハイディアン ディストリクト, シャンディ 10 ス
ストリート, ナンバー 10, バイドウ キャンパス 2 / フロア
- (72)発明者 スーイー ディーン
中華人民共和国, 100085 ベイジン, ハイディアン ディストリクト, シャンディ 10 ス
ストリート, ナンバー 10, バイドウ キャンパス 2 / フロア
- (72)発明者 シュオホワン ワーン
中華人民共和国, 100085 ベイジン, ハイディアン ディストリクト, シャンディ 10 ス
ストリート, ナンバー 10, バイドウ キャンパス 2 / フロア
- (72)発明者 ユイ スウン
中華人民共和国, 100085 ベイジン, ハイディアン ディストリクト, シャンディ 10 ス
ストリート, ナンバー 10, バイドウ キャンパス 2 / フロア
- (72)発明者 ハオ ティエン
中華人民共和国, 100085 ベイジン, ハイディアン ディストリクト, シャンディ 10 ス
ストリート, ナンバー 10, バイドウ キャンパス 2 / フロア
- (72)発明者 ホワ ウー
中華人民共和国, 100085 ベイジン, ハイディアン ディストリクト, シャンディ 10 ス
ストリート, ナンバー 10, バイドウ キャンパス 2 / フロア
- (72)発明者 ハイフオン ワーン
中華人民共和国, 100085 ベイジン, ハイディアン ディストリクト, シャンディ 10 ス
ストリート, ナンバー 10, バイドウ キャンパス 2 / フロア
- 審査官 多賀 実
- (56)参考文献 特開2023-12493 (JP, A)
特開2019-192073 (JP, A)
中国特許出願公開第113961686 (CN, A)
- (58)調査した分野 (Int.Cl., DB名)
G06N 3/00 - 99/00
G06F 16/00 - 16/958
G06F 18/00 - 18/40
G06F 40/00 - 40/58