

【公報種別】特許法第17条の2の規定による補正の掲載  
 【部門区分】第6部門第3区分  
 【発行日】令和6年2月15日(2024.2.15)

【国際公開番号】WO2021/252845  
 【公表番号】特表2023-530423(P2023-530423A)  
 【公表日】令和5年7月18日(2023.7.18)  
 【年通号数】公開公報(特許)2023-133  
 【出願番号】特願2022-576162(P2022-576162)  
 【国際特許分類】

10

G 0 6 F 4 0 / 5 6 ( 2 0 2 0 . 0 1 )

G 0 6 F 4 0 / 4 4 ( 2 0 2 0 . 0 1 )

G 0 6 F 4 0 / 2 7 9 ( 2 0 2 0 . 0 1 )

G 0 6 F 1 6 / 9 0 ( 2 0 1 9 . 0 1 )

【 F I 】

G 0 6 F 4 0 / 5 6

G 0 6 F 4 0 / 4 4

G 0 6 F 4 0 / 2 7 9

G 0 6 F 1 6 / 9 0 1 0 0

20

【手続補正書】

【提出日】令和6年2月6日(2024.2.6)

【手続補正1】

【補正対象書類名】特許請求の範囲

【補正対象項目名】全文

【補正方法】変更

【補正の内容】

【特許請求の範囲】

【請求項1】

方法であって、

データ処理システムが、エンティティをカバーするための値のリストを生成することと

30

、  
 前記データ処理システムが、データセットから、前記エンティティのコンテキストを有する発話を選択することと、

前記データ処理システムが、前記発話を複数のテンプレートに変換することとを備え、前記複数のテンプレートの各々は、前記値のリストにマッピングする前記エンティティのスロットを含み、前記方法はさらに、

前記データ処理システムが、前記複数のテンプレートからテンプレートを選択することと、

前記データ処理システムが、選択された前記テンプレート内の前記スロットと前記エンティティの前記値のリストとの間のマッピングに基づいて、前記値のリストから値を選択することと、

40

前記データ処理システムが、選択された前記テンプレートと選択された前記値とを用いて、人工発話を生成することとを備え、前記人工発話を生成することは、選択された前記値を、選択された前記テンプレートの前記スロットに挿入することを含む、方法。

【請求項2】

前記データセットから発話を選択する前に、前記データ処理システムが、訓練データセットを用いてモデルを訓練することとをさらに備え、前記訓練データセットは、前記データセットまたはそのサブセットであり、前記方法はさらに、

前記データ処理システムが、テストデータセットまたは検証データセットを用いて、前

50

記モデルをテストすることと、

前記データ処理システムが、前記モデルをテストすることに基づいて、前記モデルの性能を評価することと、

前記データ処理システムが、前記評価することに基づいて、前記エンティティを特定することとを備える、請求項 1 に記載の方法。

【請求項 3】

前記データ処理システムが、前記データセットまたは前記訓練データセットを前記人工発話で拡張して、拡張訓練データセットを生成することと、

前記データ処理システムが、前記拡張訓練データセットを用いて、前記モデルを訓練することと、

前記データ処理システムが、前記テストデータセットもしくは前記検証データセット、または他のテストデータセットもしくは他の検証データセットを用いて、前記モデルを再テストすることと、

前記データ処理システムが、前記モデルを再テストすることに基づいて、前記モデルの性能を評価することと、

前記データ処理システムが、前記モデルの前記性能が既定の規準を満たすと判断することと、

前記モデルが前記既定の規準を満たすことに応答して、前記モデルを、チャットボットシステムの一部としてデプロイすることとをさらに備える、請求項 2 に記載の方法。

【請求項 4】

前記データ処理システムが、選択された前記テンプレート内の前記スロットが、前記値のリストにマッピングする第 1 の要素と、関数にマッピングする第 2 の要素とを含むと判断することと、

前記データ処理システムが、前記関数を処理して、選択された前記テンプレート内の前記スロットと前記関数との間のマッピングに基づいて、関数値を生成することと、

前記データ処理システムが、選択された前記値と前記関数値とを組み合わせ、スロット値を取得することとをさらに備え、

前記人工発話を生成することは、前記スロット値を、選択された前記テンプレートの前記スロットに挿入することを含む、請求項 1 ~ 3 のいずれか 1 項に記載の方法。

【請求項 5】

前記データ処理システムが、選択された前記テンプレートが、前記スロットと他のスロットとを含むスロットマッピングの入れ子セットにマッピングすると判断することと、

前記データ処理システムが、前記スロットマッピングの入れ子セットの前記他のスロットが、他の値のリストにマッピングすると判断することと、

前記データ処理システムが、選択された前記テンプレート内の前記他のスロットと前記他の値のリストとの間のマッピングに基づいて、前記他の値のリストから他の値を選択することとをさらに備え、

前記人工発話を生成することは、選択された前記値を、選択された前記テンプレートの前記スロットに挿入することと、選択された前記他の値を、選択された前記テンプレートの前記他のスロットに挿入することとを含む、請求項 1 ~ 3 のいずれか 1 項に記載の方法。

【請求項 6】

前記データ処理システムが、前記データセットまたは訓練データセットを前記人工発話で拡張して、前記エンティティを含む複数のエンティティを認識および分類するようにモデルを訓練するための拡張訓練データセットを生成することをさらに備える、請求項 1 に記載の方法。

【請求項 7】

前記テンプレートは、( i ) 複数のエンティティクラスを有するテンプレートか単一のエンティティクラスを有するテンプレートかの選択、( i i ) テンプレートの多様性かテンプレートの品質かのトレードオフ、( i i i ) コンテキストに応じたテンプレートの分

10

20

30

40

50

類、(iv)ユーザがモデルに教示したいテンプレート/固有コンテキストごとの最小発生回数を前記モデルが必要とするか、または(v)これらの任意の組み合わせ、という条件を含むヒューリスティックに基づいて選択される、請求項1~6のいずれか1項に記載の方法。

【請求項8】

システムであって、

1つまたは複数のプロセッサと、

前記1つまたは複数のプロセッサに結合されたメモリとを備え、前記メモリは、前記1つまたは複数のプロセッサによって実行可能な複数の命令を格納し、前記複数の命令は、前記1つまたは複数のプロセッサによって実行されると、前記1つまたは複数のプロセッサに処理を実行させ、前記処理は、

エンティティをカバーするための値のリストを生成することと、

データセットから、前記エンティティのコンテキストを有する発話を選択することと

、  
前記発話を複数のテンプレートに変換することとを含み、前記複数のテンプレートの各々は、前記値のリストにマッピングする前記エンティティのロットを含み、前記処理はさらに、

前記複数のテンプレートからテンプレートを選択することと、

選択された前記テンプレート内の前記ロットと前記エンティティの前記値のリストとの間のマッピングに基づいて、前記値のリストから値を選択することと、

選択された前記テンプレートと選択された前記値とを用いて、人工発話を生成することとを含み、前記人工発話を生成することは、選択された前記値を、選択された前記テンプレートの前記ロットに挿入することを含む、システム。

【請求項9】

前記処理はさらに、

前記データセットから発話を選択する前に、訓練データセットを用いてモデルを訓練することを含み、前記訓練データセットは、前記データセットまたはそのサブセットであり、前記処理はさらに、

テストデータセットまたは検証データセットを用いて、前記モデルをテストすることと

、  
前記モデルをテストすることに基づいて、前記モデルの性能を評価することと、

前記評価することに基づいて、前記エンティティを特定することとを含む、請求項8に記載のシステム。

【請求項10】

前記処理はさらに、

前記データセットまたは前記訓練データセットを前記人工発話で拡張して、拡張訓練データセットを生成することと、

前記拡張訓練データセットを用いて、前記モデルを訓練することと、

前記テストデータセットもしくは前記検証データセット、または他のテストデータセットもしくは他の検証データセットを用いて、前記モデルを再テストすることと、

前記モデルを再テストすることに基づいて、前記モデルの性能を評価することと、

前記モデルの前記性能が既定の規準を満たすと判断することと、

前記モデルが前記既定の規準を満たすことに応答して、前記モデルを、チャットボットシステムの一部としてデプロイすることとを含む、請求項9に記載のシステム。

【請求項11】

前記処理はさらに、

選択された前記テンプレート内の前記ロットが、前記値のリストにマッピングする第1の要素と、関数にマッピングする第2の要素とを含むと判断することと、

前記関数を処理して、選択された前記テンプレート内の前記ロットと前記関数との間のマッピングに基づいて、関数値を生成することと、

10	前記1つまたは複数のプロセッサによって実行可能な複数の命令を格納し、前記複数の命令は、前記1つまたは複数のプロセッサによって実行されると、前記1つまたは複数のプロセッサに処理を実行させ、前記処理は、
20	エンティティをカバーするための値のリストを生成することと、
20	データセットから、前記エンティティのコンテキストを有する発話を選択することと
30	、
30	前記発話を複数のテンプレートに変換することとを含み、前記複数のテンプレートの各々は、前記値のリストにマッピングする前記エンティティのロットを含み、前記処理はさらに、
40	前記複数のテンプレートからテンプレートを選択することと、
40	選択された前記テンプレート内の前記ロットと前記エンティティの前記値のリストとの間のマッピングに基づいて、前記値のリストから値を選択することと、
40	選択された前記テンプレートと選択された前記値とを用いて、人工発話を生成することとを含み、前記人工発話を生成することは、選択された前記値を、選択された前記テンプレートの前記ロットに挿入することを含む、システム。
50	【請求項9】
50	前記処理はさらに、
50	前記データセットから発話を選択する前に、訓練データセットを用いてモデルを訓練することを含み、前記訓練データセットは、前記データセットまたはそのサブセットであり、前記処理はさらに、
50	テストデータセットまたは検証データセットを用いて、前記モデルをテストすることと
50	、
50	前記モデルをテストすることに基づいて、前記モデルの性能を評価することと、
50	前記評価することに基づいて、前記エンティティを特定することとを含む、請求項8に記載のシステム。
50	【請求項10】
50	前記処理はさらに、
50	前記データセットまたは前記訓練データセットを前記人工発話で拡張して、拡張訓練データセットを生成することと、
50	前記拡張訓練データセットを用いて、前記モデルを訓練することと、
50	前記テストデータセットもしくは前記検証データセット、または他のテストデータセットもしくは他の検証データセットを用いて、前記モデルを再テストすることと、
50	前記モデルを再テストすることに基づいて、前記モデルの性能を評価することと、
50	前記モデルの前記性能が既定の規準を満たすと判断することと、
50	前記モデルが前記既定の規準を満たすことに応答して、前記モデルを、チャットボットシステムの一部としてデプロイすることとを含む、請求項9に記載のシステム。
50	【請求項11】
50	前記処理はさらに、
50	選択された前記テンプレート内の前記ロットが、前記値のリストにマッピングする第1の要素と、関数にマッピングする第2の要素とを含むと判断することと、
50	前記関数を処理して、選択された前記テンプレート内の前記ロットと前記関数との間のマッピングに基づいて、関数値を生成することと、

選択された前記値と前記関数値とを組み合わせ、スロット値を取得することとをさらに含み、

前記人工発話を生成することは、前記スロット値を、選択された前記テンプレートの前記スロットに挿入することを含む、請求項 8 ~ 10 のいずれか 1 項に記載のシステム。

【請求項 1 2】

前記処理はさらに、

選択された前記テンプレートが、前記スロットと他のスロットとを含むスロットマッピングの入れ子セットにマッピングすると判断することと、

前記スロットマッピングの入れ子セットの前記他のスロットが、他の値のリストにマッピングすると判断することと、

選択された前記テンプレート内の前記他のスロットと前記他の値のリストとの間のマッピングに基づいて、前記他の値のリストから他の値を選択することとを含み、

前記人工発話を生成することは、選択された前記値を、選択された前記テンプレートの前記スロットに挿入することと、選択された前記他の値を、選択された前記テンプレートの前記他のスロットに挿入することを含む、請求項 8 ~ 10 のいずれか 1 項に記載のシステム。

【請求項 1 3】

前記処理はさらに、前記データセットまたは訓練データセットを前記人工発話で拡張して、前記エンティティを含む複数のエンティティを認識および分類するようにモデルを訓練するための拡張訓練データセットを生成することを含む、請求項 8 に記載のシステム。

【請求項 1 4】

前記テンプレートは、( i ) 複数のエンティティクラスを有するテンプレートか単一のエンティティクラスを有するテンプレートかの選択、( i i ) テンプレートの多様性かテンプレートの品質かのトレードオフ、( i i i ) コンテキストに応じたテンプレートの分類、( i v ) ユーザがモデルに教示したいテンプレート / 固有コンテキストごとの最小発生回数を前記モデルが必要とするか、または ( v ) これらの任意の組み合わせ、という条件を含むヒューリスティックに基づいて選択される、請求項 8 ~ 1 3 のいずれか 1 項に記載のシステム。

【請求項 1 5】

請求項 1 ~ 7 のいずれか 1 項に記載の方法を 1 つまたは複数のプロセッサに実行させるためのコンピュータ読取可能プログラム。

10

20

30

40

50