【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8月1日(2024.8.1)

【国際公開番号】WO2022/046759
【公表番号】特表2023-539240(P2023-539240A)
【公表日】令和5年9月13日(2023.9.13)
【年通号数】公開公報(特許)2023-173
【出願番号】特願2023-513274(P2023-513274)
【国際特許分類】
Ｇ０６Ｎ 20/00(2019.01)
Ｇ０６Ｆ 8/76(2018.01)
【ＦＩ】
Ｇ０６Ｎ 20/00
Ｇ０６Ｆ 8/76

【手続補正書】
【提出日】令和6年7月24日(2024.7.24)
【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】
データセットを受信するステップと、
前記受信したデータセットに対して自動機械学習プロセスを実行して、新しいデータに基づいて予測を行うためのモデルをトレーニングするステップであって、前記自動機械学習プロセスは、パイプラインに基づいて一連の決定を行うステップを含む、ステップと、
前記モデルがトレーニングされた後、前記パイプラインの表現を生成するステップと、
ユーザインターフェースにおいて表示するための前記パイプラインの前記表現を提供するステップであって、前記ユーザインターフェースは、ユーザが前記自動機械学習プロセスにおいて行われた前記決定の少なくとも一部を修正するステップを可能にする、ステップと、
前記ユーザインターフェースを介して前記ユーザから１つまたは複数の修正を受信するステップと、
前記ユーザからの前記１つまたは複数の修正に基づいて前記モデルを改良するステップであって、前記改良されたモデルは、新しいデータに基づいて前記予測を行うために使用される、ステップと、
を含むコンピュータ実装方法。
【請求項２】
各決定は、前記自動機械学習プロセス中に前記パイプライン内の一連のステップのステップにおいて行われ、前記パイプラインの前記表現を生成するステップは、
前記一連のステップの１つまたは複数のステップのそれぞれについての表現を生成するステップであって、ステップの前記表現は、前記ステップについての複数のオプションを含み、前記複数のオプションは、前記自動機械学習プロセス中に前記ステップにおいて行われた前記決定を含む、ステップを含む、請求項１に記載の方法。
【請求項３】
前記ステップの前記表現は、前記複数のオプションのそれぞれについてのランキングスコアをさらに含み、オプションの前記ランキングスコアは、前記オプションに対する推奨

事項のレベルを示す、請求項２に記載の方法。
【請求項４】
前記パイプラインの前記表現を生成するステップは、
前記モデルをトレーニングするために前記自動機械学習プロセスにおいて使用される特徴、前記特徴についての複数のオプションのデータタイプ、および各オプションのデータタイプについてのランキングスコアを含むデータタイプリストを生成するステップであって、前記複数のオプションのデータタイプは、前記モデルをトレーニングするための前記自動機械学習プロセスにおける前記特徴のために選択された第１のデータタイプを含み、オプションのデータタイプの前記ランキングスコアは、前記データタイプが前記特徴の真のデータタイプである確率を示す、ステップを含む、請求項１に記載の方法。
【請求項５】
前記ユーザインターフェースを介して前記ユーザから１つまたは複数の修正を受信するステップは、
前記ユーザインターフェースを介して前記ユーザから、前記複数のオプションのデータタイプからの前記特徴についての第２のデータタイプの選択を受信するステップであって、前記第２のデータタイプは、前記第１のデータタイプとは異なるデータタイプである、ステップと、
前記第２のデータタイプに基づいて前記特徴の値を符号化するステップと、
を含む、請求項４に記載の方法。
【請求項６】
前記パイプラインの前記表現を生成するステップは、
特徴リストを生成するステップであって、前記特徴リストは、複数の特徴と、前記ユーザが前記予測にとってどの程度重要であるかを示す各特徴の説明と、を含み、前記複数の特徴は、前記モデルをトレーニングするために前記自動機械学習プロセスにおいて使用される特徴を含む、ステップを含む、請求項１に記載の方法。
【請求項７】
前記パイプラインの前記表現を生成するステップは、
アルゴリズムリストを生成するステップであって、前記アルゴリズムリストは、複数のアルゴリズムと、前記モデルをトレーニングするために前記アルゴリズムを選択するまたは選択しない理由を示す各アルゴリズムの説明とを含み、前記複数のアルゴリズムは、前記モデルをトレーニングするために前記自動機械学習プロセスにおいて使用されるアルゴリズムを含む、ステップを含む、請求項１に記載の方法。
【請求項８】
データ分析システムにおいてデータブロックを処理するための実行可能なコンピュータプログラム命令を格納する非一時的コンピュータ可読メモリであって、前記コンピュータプログラム命令は
データセットを受信するステップと、
前記受信したデータセットに対して自動機械学習プロセスを実行して、新しいデータに基づいて予測を行うためのモデルをトレーニングするステップであって、前記自動機械学習プロセスは、パイプラインに基づいて一連の決定を行うステップを含む、ステップと、
前記モデルがトレーニングされた後、前記パイプラインの表現を生成するステップと、
ユーザインターフェースにおいて表示するための前記パイプラインの前記表現を提供するステップであって、前記ユーザインターフェースは、ユーザが前記自動機械学習プロセスにおいて行われた前記決定の少なくとも一部を修正するステップを可能にする、ステップと、
前記ユーザインターフェースを介して前記ユーザから１つまたは複数の修正を受信するステップと、
前記ユーザからの前記１つまたは複数の修正に基づいて前記モデルを改良するステップであって、前記改良されたモデルは、新しいデータに基づいて前記予測を行うために使用される、ステップと、

を含む動作を実行するステップを実行可能である、非一時的コンピュータ可読メモリ。
【請求項９】
各決定は、前記自動機械学習プロセス中に前記パイプライン内の一連のステップのステップにおいて行われ、前記パイプラインの前記表現を生成するステップは、
前記一連のステップの１つまたは複数のステップのそれぞれについての表現を生成するステップであって、ステップの前記表現は、前記ステップについての複数のオプションを含み、前記複数のオプションは、前記自動機械学習プロセス中に前記ステップにおいて行われた前記決定を含む、ステップを含む、請求項８に記載の非一時的コンピュータ可読メモリ。
【請求項１０】
前記ステップの前記表現は、前記複数のオプションのそれぞれについてのランキングスコアをさらに含み、オプションの前記ランキングスコアは、前記オプションに対する推奨事項のレベルを示す、請求項９に記載の非一時的コンピュータ可読メモリ。
【請求項１１】
前記パイプラインの前記表現を生成するステップは、
前記モデルをトレーニングするために前記自動機械学習プロセスにおいて使用される特徴、前記特徴についての複数のオプションのデータタイプ、および各オプションのデータタイプについてのランキングスコアを含むデータタイプリストを生成するステップであって、前記複数のオプションのデータタイプは、前記モデルをトレーニングするための前記自動機械学習プロセスにおける前記特徴のために選択された第１のデータタイプを含み、オプションのデータタイプの前記ランキングスコアは、前記データタイプが前記特徴の真のデータタイプである確率を示す、ステップを含む、請求項８に記載の非一時的コンピュータ可読メモリ。
【請求項１２】
前記ユーザインターフェースを介して前記ユーザから１つまたは複数の修正を受信するステップは、
前記ユーザインターフェースを介して前記ユーザから、前記複数のオプションのデータタイプからの前記特徴についての第２のデータタイプの選択を受信するステップであって、前記第２のデータタイプは、前記第１のデータタイプとは異なるデータタイプである、ステップと、
前記第２のデータタイプに基づいて前記特徴の値を符号化するステップと、
を含む、請求項１１に記載の非一時的コンピュータ可読メモリ。
【請求項１３】
前記パイプラインの前記表現を生成するステップは、
特徴リストを生成するステップであって、前記特徴リストは、複数の特徴と、前記ユーザが前記予測にとってどの程度重要であるかを示す各特徴の説明と、を含み、前記複数の特徴は、前記モデルをトレーニングするために前記自動機械学習プロセスにおいて使用される特徴を含む、ステップを含む、請求項８に記載の非一時的コンピュータ可読メモリ。
【請求項１４】
前記パイプラインの前記表現を生成するステップは、
アルゴリズムリストを生成するステップであって、前記アルゴリズムリストは、複数のアルゴリズムと、前記モデルをトレーニングするために前記アルゴリズムを選択するまたは選択しない理由を示す各アルゴリズムの説明とを含み、前記複数のアルゴリズムは、前記モデルをトレーニングするために前記自動機械学習プロセスにおいて使用されるアルゴリズムを含む、ステップ、を含む、請求項８に記載の非一時的コンピュータ可読メモリ。
【請求項１５】
コンピュータプログラム命令を実行するためのコンピュータプロセッサと、
動作を実行する前記コンピュータプロセッサによって実行可能なコンピュータプログラム命令を格納する非一時的コンピュータ可読メモリと、を含むシステムであって、前記動作は、

データセットを受信するステップと、
前記受信したデータセットに対して自動機械学習プロセスを実行して、新しいデータに基づいて予測を行うためのモデルをトレーニングするステップであって、前記自動機械学習プロセスは、パイプラインに基づいて一連の決定を行うステップを含む、ステップと、
前記モデルがトレーニングされた後、前記パイプラインの表現を生成するステップと、
ユーザインターフェースにおいて表示するための前記パイプラインの前記表現を提供するステップであって、前記ユーザインターフェースは、ユーザが前記自動機械学習プロセスにおいて行われた前記決定の少なくとも一部を修正するステップを可能にする、ステップと、
前記ユーザインターフェースを介して前記ユーザから１つまたは複数の修正を受信するステップと、
前記ユーザからの前記１つまたは複数の修正に基づいて前記モデルを改良するステップであって、前記改良されたモデルは、新しいデータに基づいて前記予測を行うために使用される、ステップと、
を含むシステム。
【請求項１６】
各決定は、前記自動機械学習プロセス中に前記パイプライン内の一連のステップのステップにおいて行われ、前記パイプラインの前記表現を生成するステップは、
前記一連のステップの１つまたは複数のステップのそれぞれについての表現を生成するステップであって、ステップの前記表現は、前記ステップについての複数のオプションを含み、前記複数のオプションは、前記自動機械学習プロセス中に前記ステップにおいて行われた前記決定を含む、ステップを含む、請求項１５に記載のシステム。
【請求項１７】
前記ステップの前記表現は、前記複数のオプションのそれぞれについてのランキングスコアをさらに含み、オプションの前記ランキングスコアは、前記オプションに対する推奨事項のレベルを示す、請求項１６に記載のシステム。
【請求項１８】
前記パイプラインの前記表現を生成するステップは、
前記モデルをトレーニングするために前記自動機械学習プロセスにおいて使用される特徴、前記特徴についての複数のオプションのデータタイプ、および各オプションのデータタイプについてのランキングスコアを含むデータタイプリストを生成するステップであって、前記複数のオプションのデータタイプは、前記モデルをトレーニングするための前記自動機械学習プロセスにおける前記特徴のために選択された第１のデータタイプを含み、オプションのデータタイプの前記ランキングスコアは、前記データタイプが前記特徴の真のデータタイプである確率を示す、ステップを含む、請求項１５に記載のシステム。
【請求項１９】
前記パイプラインの前記表現を生成するステップは、
特徴リストを生成するステップであって、前記特徴リストは、複数の特徴と、前記ユーザが前記予測にとってどの程度重要であるかを示す各特徴の説明と、を含み、前記複数の特徴は、前記モデルをトレーニングするために前記自動機械学習プロセスにおいて使用される特徴を含む、ステップを含む、請求項１５に記載のシステム。
【請求項２０】
前記パイプラインの前記表現を生成するステップは、
アルゴリズムリストを生成するステップであって、前記アルゴリズムリストは、複数のアルゴリズムと、前記モデルをトレーニングするために前記アルゴリズムを選択するまたは選択しない理由を示す各アルゴリズムの説明とを含み、前記複数のアルゴリズムは、前記モデルをトレーニングするために前記自動機械学習プロセスにおいて使用されるアルゴリズムを含む、ステップを含む、請求項１５に記載のシステム。