【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22日(2024.2.22)

【公開番号】特開2023-37242(P2023-37242A)
【公開日】令和5年3月15日(2023.3.15)
【年通号数】公開公報(特許)2023-049
【出願番号】特願2021-143871(P2021-143871)
【国際特許分類】
Ｇ０６Ｎ 20/00(2019.01)
【ＦＩ】
Ｇ０６Ｎ 20/00

【手続補正書】
【提出日】令和6年2月14日(2024.2.14)
【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】
データ管理システムであって、
１以上の記憶装置と、
１以上の演算装置と、を含み、
前記１以上の記憶装置は、入力データ及び第１学習モデルを格納し、
前記入力データは、複数のレコードを含み、
前記複数のレコードの各レコードは、複数の特徴量の値を含み、
前記１以上の演算装置は、
前記複数の特徴量から、指定されている人関連ＩＤと異なる特徴量を選択し、
前記選択された特徴量のデータによって、第１学習モデルの訓練を実行し、
訓練済みの前記第１学習モデルの推定結果に基づき、前記選択した特徴量の少なくとも一部の、前記第１学習モデルの推定に対する影響度を算出し、
前記影響度及び前記少なくとも一部の特徴量の情報を、ユーザに提示する、データ管理システム。
【請求項２】
請求項１に記載のデータ管理システムであって、
前記１以上の演算装置は、
前記選択された特徴量のデータの間の相関を決定し、
他の特徴量との相関が閾値を超える特徴量から、前記影響度の算出から除外する特徴量を決定する、データ管理システム。
【請求項３】
請求項１に記載のデータ管理システムであって、
前記１以上の演算装置は、
前記選択された特徴量と、指定されているセンシティブ特徴量との間の相関を決定し、
前記センシティブ特徴量との間の相関が閾値を超える特徴量から、前記影響度の算出から除外する特徴量を決定する、データ管理システム。
【請求項４】
請求項１に記載のデータ管理システムであって、
前記１以上の演算装置は、

前記選択された特徴量と、指定されているセンシティブ特徴量との間の相関を決定し、
前記センシティブ特徴量との間の相関が閾値未満の特徴量から、前記影響度の算出から除外する特徴量を決定する、データ管理システム。
【請求項５】
請求項１に記載のデータ管理システムであって、
前記１以上の記憶装置は、前記入力データによる訓練の目的とする第２学習モデルを格納し、
前記第１学習モデルの訓練の処理負荷は、前記第２学習モデルの訓練の処理負荷より小さい、データ管理システム。
【請求項６】
請求項５に記載のデータ管理システムであって、
前記１以上の演算装置は、
前記影響度及び前記少なくとも一部の特徴量の情報を前記ユーザに提示した後、前記入力データにおいて前記ユーザに選択された特徴量のデータによって、前記第２学習モデルを訓練し、
訓練された前記第２学習モデルの推定結果に基づき、前記第２学習モデルの公平性及び／又は正確性の指標を決定して、前記ユーザに提示する、データ管理システム。
【請求項７】
請求項１に記載のデータ管理システムであって、
前記１以上の演算装置は、前記第１学習モデルの推定結果から得られる混同行列に基づいて、前記影響度を算出する、データ管理システム。
【請求項８】
請求項１に記載のデータ管理システムであって、
前記１以上の記憶装置は、過去の分析処理履歴を格納し、
前記１以上の演算装置は、前記選択された特徴量の情報を前記分析処理履歴において検索し、
前記選択された特徴量と同一特徴量の情報が前記分析処理履歴において見つかると、前記分析処理履歴から取得した影響度を前記ユーザに提示する、データ管理システム。
【請求項９】
システムがデータを管理する、管理方法あって、
前記システムは入力データ及び第１学習モデルを格納し、
前記入力データは、複数のレコードを含み、
前記複数のレコードの各レコードは、複数の特徴量の値を含み、
前記管理方法は、前記システムが、
前記複数の特徴量から、指定されている人関連ＩＤと異なる特徴量を選択し、
前記選択された特徴量のデータによって、第１学習モデルの訓練を実行し、
訓練済みの前記第１学習モデルの推定結果に基づき、前記選択した特徴量の少なくとも一部の、前記第１学習モデルの推定に対する影響度を算出し、
前記影響度及び前記少なくとも一部の特徴量の情報を、ユーザに提示する、ことを含む、管理方法。