

【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 3 年 5 月 6 日 (2021.5.6)

【公表番号】特表 2021-509502 (P2021-509502A)
【公表日】令和 3 年 3 月 25 日 (2021.3.25)
【年通号数】公開・登録公報 2021-015
【出願番号】特願 2020-535049 (P2020-535049)
【国際特許分類】

G 1 6 H 50/00 (2018.01)

G 1 6 B 40/30 (2019.01)

【F I】

G 1 6 H 50/00

G 1 6 B 40/30

【手続補正書】
【提出日】令和 3 年 2 月 19 日 (2021.2.19)
【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】

コンピュータ実装方法であって、

1 つまたは複数のプロセッサにより、ある状態と診断もしくは予備診断されるか、または、前記状態と診断されないものとして指定される患者のペプチド・ライブラリを含む生データ・セットを取得することと、

前記 1 つまたは複数のプロセッサにより、前記生データ・セットを所定数のグループに分割することであり、不適合データ・グループをその他のグループから取り出すことを含む、前記分割することと、

前記 1 つまたは複数のプロセッサにより、前記その他のグループに対して主成分分析を実行して、前記その他のグループにおける特徴の頻度に基づいて、前記 1 つまたは複数のプロセッサにより、前記その他のグループのデータにおける共通特徴を識別し、前記その他のグループにおける発生頻度に基づいて、係数により表される主成分を含む前記共通特徴を重み付けすることと、

前記 1 つまたは複数のプロセッサにより、所定レベルの検証精度を生じる最小数の前記主成分を決定することと、

前記 1 つまたは複数のプロセッサにより、ロジスティック回帰モデルにおける最良適合のパラメータとして前記最小数の主成分を利用することによって予測モデルを生成することであり、前記予測モデルが、所定の閾値内に前記状態が存在する尤度および前記所定の閾値内に前記状態が存在しない尤度から成る群から選択される 2 値結果を提供する、前記生成することと、

を含むコンピュータ実装方法。

【請求項 2】

前記 1 つまたは複数のプロセッサにより、前記係数を適用して前記不適合データ・グループに使用する特徴を計算することによって前記予測モデルをテストすることと、

前記 1 つまたは複数のプロセッサにより、前記不適合データ・グループの前記特徴を前記主成分と比較することによって前記予測モデルを調節することと、

をさらに含む、請求項 1 に記載のコンピュータ実装方法。

【請求項 3】

前記ペプチド・ライブラリが、12-mer ペプチド・ライブラリを含む、請求項 1 に記載のコンピュータ実装方法。

【請求項 4】

前記状態が、ステージ 1 の乳がんを含む、請求項 1 に記載のコンピュータ実装方法。

【請求項 5】

前記分割が、前記グループに割り当てる前記生データ・セットのレコードを選択する乱数を生成することを含む、請求項 1 に記載のコンピュータ実装方法。

【請求項 6】

前記主成分分析の実行に先立って、前記 1 つまたは複数のプロセッサにより、すべてのペプチドにわたりペプチドのペプチド値を合算するとともに各発現値を結果の合計値で割って比を演算することによって、前記生データ・セットを含むペプチドを正規化することを含む、請求項 1 に記載のコンピュータ実装方法。

【請求項 7】

前記その他のグループに対して主成分分析を実行することが、

前記 1 つまたは複数のプロセッサにより、トレーニング・データ集合を生成することであり、前記トレーニング・データ集合の各トレーニング・データ集合が、前記その他のグループのうちの 1 つを除くすべてのグループのデータを含む、前記生成することと、

前記 1 つまたは複数のプロセッサにより、各トレーニング・データ集合に対して個別に主成分分析を実行することによって、各トレーニング・データ集合の共通特徴を識別することと、

前記 1 つまたは複数のプロセッサにより、前記その他のグループを単一のトレーニング・セットとして利用することにより主成分分析を実行することによって、前記単一のトレーニング・セットの共通特徴を識別することと、

各トレーニング・データ集合の前記共通特徴および前記単一のトレーニング・セットの前記共通特徴について、前記その他のグループの数に等しい回数の相互検証を実行することによって、前記主成分を識別するとともに前記係数を決定することと、

を含む、請求項 1 に記載のコンピュータ実装方法。

【請求項 8】

前記所定レベルの検証精度が、前記回数に基づく最高の検証精度である、請求項 7 に記載のコンピュータ実装方法。

【請求項 9】

前記 1 つまたは複数のプロセッサにより、所与の個人の 12-mer ペプチド・ライブラリを含む新たな生データを取得することと、

前記 1 つまたは複数のプロセッサにより、前記予測モデルを前記新たな生データに適用して、前記所与の個人の 2 値結果を決定することと、

をさらに含む、請求項 1 に記載のコンピュータ実装方法。

【請求項 10】

前記予測モデルを適用することが、前記係数を利用して前記新たな生データに対する主成分分析を実行することと、ロジスティック回帰を実行して前記 2 値結果を生成することと、を含む、請求項 9 に記載のコンピュータ実装方法。

【請求項 11】

コンピュータ・プログラムであって、請求項 1 ないし 10 のいずれか 1 項に記載の方法の各ステップをコンピュータに実行させるための、コンピュータ・プログラム。

【請求項 12】

請求項 11 に記載のコンピュータ・プログラムを記録した、コンピュータ可読記憶媒体。

【請求項 13】

システムであって、
メモリと、

前記メモリと通信するプロセッサと、

方法を実行するように前記メモリを介して前記プロセッサにより実行可能なプログラム命令と、

を備え、前記方法は、

前記 1 つまたは複数のプロセッサにより、ある状態と診断もしくは予備診断されるか、または、前記状態と診断されないものとして指定される患者のペプチド・ライブラリを含む生データ・セットを取得することと、

前記 1 つまたは複数のプロセッサにより、前記生データ・セットを所定数のグループに分割することであり、不適合データ・グループをその他のグループから取り出すことを含む、前記分割することと、

前記 1 つまたは複数のプロセッサにより、前記その他のグループに対して主成分分析を実行して、前記その他のグループにおける特徴の頻度に基づいて、前記 1 つまたは複数のプロセッサにより、前記その他のグループのデータにおける共通特徴を識別し、その他のグループにおける発生頻度に基づいて、係数により表される主成分を含む前記共通特徴を重み付けすることと、

前記 1 つまたは複数のプロセッサにより、所定レベルの検証精度を生じる最小数の前記主成分を決定することと、

前記 1 つまたは複数のプロセッサにより、ロジスティック回帰モデルにおける最良適合のパラメータとして前記最小数の主成分を利用することによって予測モデルを生成することであり、前記予測モデルが、所定の閾値内に前記状態が存在する尤度および前記所定の閾値内に前記状態が存在しない尤度から成る群から選択される 2 値結果を提供する、前記生成することと、

を含む、システム。