(12) 特許協力条約に基づいて公開された国際出願

(19) 世界知的所有権機関
国際事務局

(43) 国際公開日
2025年4月3日(03.04.2025)

WIPO | PCT

(10) 国際公開番号
WO 2025/070104 A1

(51) 国際特許分類:
G16B 30/00 (2019.01)　*G16B 40/00* (2019.01)

(21) 国際出願番号：　PCT/JP2024/032771

(22) 国際出願日：　2024年9月12日(12.09.2024)

(25) 国際出願の言語：　日本語

(26) 国際公開の言語：　日本語

(30) 優先権データ：
特願 2023-167582　2023年9月28日(28.09.2023)　JP

(71) 出願人: 富士フイルム株式会社 (**FUJIFILM CORPORATION**) [JP/JP]; 〒1068620 東京都港区西麻布２丁目２６番３０号 (JP).

(72) 発明者: 蔦岡　拓也 (**TSUTAOKA, Takuya**); 〒2588577 神奈川県足柄上郡開成町牛島５７７番地 富士フイルム株式会社内 (JP).

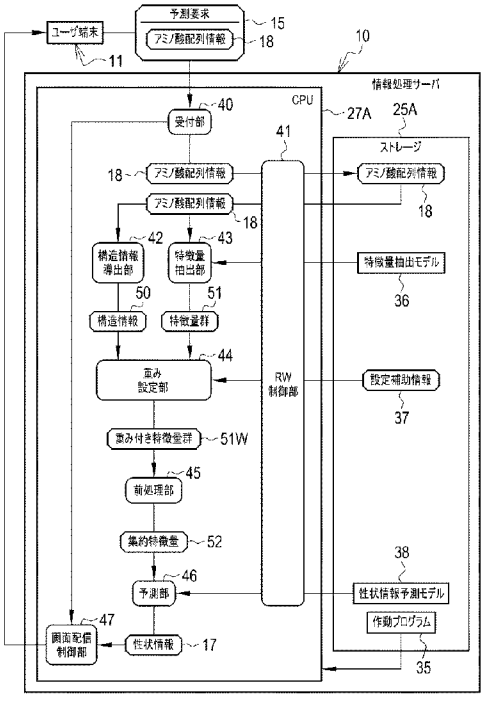
(74) 代理人: 弁理士法人太陽国際特許事務所(**TAIYO, NAKAJIMA & KATO**); 〒1600022 東京都新宿区新宿４丁目３番１７号 (JP).

(81) 指定国(表示のない限り、全ての種類の国内保護が可能): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CV, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, MG, MK, MN, MU, MW, MX, MY,

WO 2025/070104 A1

(54) **Title:** INFORMATION PROCESSING DEVICE, OPERATION METHOD FOR INFORMATION PROCESSING DEVICE, AND OPERATION PROGRAM FOR INFORMATION PROCESSING DEVICE

(54) 発明の名称： 情報処理装置、情報処理装置の作動方法、および情報処理装置の作動プログラム



10 Information processing server
11 User terminal
15 Prediction request
17 Property information
18 Amino acid sequence information
25A Storage
35 Operation program
36 Feature amount extraction model
37 Setting assistance information
38 Property information prediction model
40 Reception unit
41 RW control unit
42 Structure information derivation unit
43 Feature amount extraction unit
44 Weight setting unit
45 Preprocessing unit
46 Prediction unit
47 Screen distribution control unit
50 Structure information
51 Feature amount group
51W Weighted feature amount group
52 Integrated feature amount

(57) **Abstract:** Provided is an information processing device comprising a processor, wherein the processor: acquires structure information about an amino acid sequence unique to a protein, and a feature amount of an amino acid residue constituting the amino acid sequence; sets a weight to the feature amount on the basis of the structure information; and causes a machine-learning model to predict property information, which indicates the property of the protein, by applying,

[続葉有]

MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW.

(84) 指定国(表示のない限り、全ての種類の広域保護が可能):ARIPO (BW, CV, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SC, SD, SL, ST, SZ, TZ, UG, ZM, ZW), ユーラシア (AM, AZ, BY, KG, KZ, RU, TJ, TM), ヨーロッパ (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, ME, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

添付公開書類:

一　国際調査報告(条約第21条(3))

to the machine learning model, the feature amount to which the weight is set.

(57) 要約:プロセッサを備え、プロセッサは、タンパク質に固有のアミノ酸配列の構造情報と、アミノ酸配列を構成するアミノ酸残基の特徴量とを取得し、構造情報に基づいて特徴量の重みを設定し、重みが設定された特徴量を機械学習モデルに適用することで、機械学習モデルにタンパク質の性状を示す性状情報を予測させる、情報処理装置。

明　細　書

発明の名称 ：

情報処理装置、情報処理装置の作動方法、および情報処理装置の作動プログラム

技術分野

[0001] 本開示の技術は、情報処理装置、情報処理装置の作動方法、および情報処理装置の作動プログラムに関する。

[0002] 最近、高い薬効および副作用の少なさから、バイオ医薬品、ペプチド医薬品、核酸医薬品といった医薬品が注目を浴びている。一例として、バイオ医薬品は、例えばインターフェロン、抗体といったタンパク質を有効成分とする。こうした抗体等のタンパク質は、例えば、最も高い活性が得られる、または最も品質が安定的に維持される水素イオン指数（ｐＨ（Ｐｏｔｅｎｔｉａｌ　Ｏｆ　Ｈｙｄｒｏｇｅｎ）値）といった、その性状を示す性状情報が種類毎に異なる。

[0003] 特開２０２１－１５８９７３号公報には、アミノ酸配列においてアミノ酸置換変異を備えているタンパク質の変異体の熱的安定性（耐熱温度が設定温度よりも高いか否か等）を、機械学習モデルにより予測する技術が記載されている。特開２０２１－１５８９７３号公報では、熱的安定性の予測精度を高めるために、アミノ酸置換変異の部分を中心とするタンパク質の変異体の局所領域のアミノ酸残基の特徴量（局所領域のアミノ酸残基の種類および数、局所領域のアミノ酸残基の大きさ、側鎖の角度、疎水度等）を、選択的に機械学習モデルに適用している。

発明の概要

発明が解決しようとする課題

[0004] 特開２０２１－１５８９７３号公報に記載の技術は、タンパク質の変異体を対象としている。このため、変異体ではない一般的なタンパク質には用いることができない。

[0005] 本開示の技術に係る１つの実施形態は、タンパク質の性状を示す性状情報を高い精度で予測することが可能な情報処理装置、情報処理装置の作動方法、および情報処理装置の作動プログラムを提供する。

課題を解決するための手段

[0006] 本開示の情報処理装置は、プロセッサを備え、プロセッサは、タンパク質に固有のアミノ酸配列の構造情報と、アミノ酸配列を構成するアミノ酸残基の特徴量とを取得し、構造情報に基づいて特徴量の重みを設定し、重みが設定された特徴量を機械学習モデルに適用することで、機械学習モデルにタンパク質の性状を示す性状情報を予測させる。

[0007] 構造情報は、アミノ酸残基の機能的な区分領域を示す区分領域情報であることが好ましい。

[0008] タンパク質は抗体であり、区分領域は相補性決定領域を含み、プロセッサは、区分領域情報で示される区分領域が相補性決定領域であるアミノ酸残基の特徴量の重みを、区分領域情報で示される区分領域が相補性決定領域でないアミノ酸残基の特徴量の重みよりも高く設定することが好ましい。

[0009] 構造情報は、アミノ酸配列の立体構造情報であることが好ましい。

[0010] 立体構造情報は、アミノ酸配列の３次元的な形状を示す３次元形状情報であることが好ましい。

[0011] ３次元的な形状は、ループ形状および直線形状を含み、プロセッサは、３次元形状情報で示される３次元的な形状がループ形状であるアミノ酸残基の特徴量の重みを、３次元形状情報で示される３次元的な形状が直線形状であるアミノ酸残基の特徴量の重みよりも高く設定することが好ましい。

[0012] プロセッサは、アミノ酸残基の特徴量、または機械学習モデルに対して、重みに応じた前処理を施すことが好ましい。

[0013] 前処理は、アミノ酸残基の特徴量の重みによる加重平均を算出する処理であることが好ましい。

[0014] 前処理は、重みに応じた初期係数を機械学習モデルに設定する処理であることが好ましい。

[0015] 性状情報は、タンパク質の至適な水素イオン指数であることが好ましい。

[0016] タンパク質は抗体であることが好ましい。

[0017] 本開示の情報処理装置の作動方法は、タンパク質に固有のアミノ酸配列の構造情報と、アミノ酸配列を構成するアミノ酸残基の特徴量とを取得すること、構造情報に基づいて特徴量の重みを設定すること、並びに、重みが設定された特徴量を機械学習モデルに適用することで、機械学習モデルにタンパク質の性状を示す性状情報を予測させること、を含む。

[0018] 本開示の情報処理装置の作動プログラムは、タンパク質に固有のアミノ酸配列の構造情報と、アミノ酸配列を構成するアミノ酸残基の特徴量とを取得すること、構造情報に基づいて特徴量の重みを設定すること、並びに、重みが設定された特徴量を機械学習モデルに適用することで、機械学習モデルにタンパク質の性状を示す性状情報を予測させること、を含む処理をコンピュータに実行させる。

発明の効果

[0019] 本開示の技術によれば、タンパク質の性状を示す性状情報を高い精度で予測することが可能な情報処理装置、情報処理装置の作動方法、および情報処理装置の作動プログラムを提供することができる。

図面の簡単な説明

[0020] [図1]情報処理サーバおよびユーザ端末を示す図である。

[図2]抗体の基本構造を示す図である。

[図3]情報処理サーバおよびユーザ端末を構成するコンピュータを示すブロック図である。

[図4]情報処理サーバのＣＰＵの処理部を示すブロック図である。

[図5]構造情報を示す図である。

[図6]特徴量抽出部の処理、および特徴量群を示す図である。

[図7]設定補助情報を示す図である。

[図8]重み設定部の処理を示す図である。

[図9]重み付き特徴量群を示す図である。

[図10]集約特徴量を示す図である。

[図11]予測部の処理を示す図である。

[図12]性状情報予測モデルを構成するニューラルネットワークを示す図である。

[図13]ユーザ端末のＣＰＵの処理部を示すブロック図である。

[図14]情報入力画面を示す図である。

[図15]予測結果表示画面を示す図である。

[図16]情報処理サーバの処理手順を示すフローチャートである。

[図17]構造情報の別の例を示す図である。

[図18]設定補助情報の別の例を示す図である。

[図19]構造情報のさらに別の例を示す図である。

[図20]設定補助情報のさらに別の例を示す図である。

[図21]第２実施形態の前処理部の処理を示す図である。

発明を実施するための形態

[0021] ［第１実施形態］

一例として図１に示すように、情報処理サーバ１０は、ユーザ端末１１にネットワーク１２を介して接続されている。情報処理サーバ１０は、本開示の技術に係る「情報処理装置」の一例である。

[0022] ユーザ端末１１は、バイオ医薬品を開発する製薬会社、または製薬会社からバイオ医薬品の開発業務を受託する機関、すなわち医薬品開発業務受託機関（ＣＲＯ：Ｃｏｎｔｒａｃｔ　Ｒｅｓｅａｒｃｈ　Ｏｒｇａｎｉｚａｔｉｏｎ）に設置される。ユーザ端末１１は、製薬会社または医薬品開発業務受託機関（以下、まとめて製薬施設と表記する）においてバイオ医薬品の開発に携わるユーザＵにより操作される。ネットワーク１２は、例えばインターネットあるいは公衆通信網等のＷＡＮ（Ｗｉｄｅ　Ａｒｅａ　Ｎｅｔｗｏｒｋ）である。なお、図１においては１台のユーザ端末１１だけが情報処理サーバ１０に接続されているが、実際には複数の製薬施設の複数台のユーザ端末１１が情報処理サーバ１０に接続されている。

[0023] ユーザ端末１１は、情報処理サーバ１０に予測要求１５を送信する。予測要求１５は、バイオ医薬品の有効成分である抗体１６の性状を示す性状情報１７の予測を、情報処理サーバ１０に行わせるための要求である。予測要求１５は、抗体１６のアミノ酸配列情報１８を含む。アミノ酸配列情報１８は実験により同定され、ユーザＵがユーザ端末１１の入力デバイス３０Ｂ（図１３参照）を操作することで入力される。抗体１６は、本開示の技術に係る「タンパク質」の一例である。なお、図示は省略するが、予測要求１５は、予測要求１５の送信元のユーザ端末１１を一意に識別するための端末ＩＤ（Ｉｄｅｎｔｉｆｉｃａｔｉｏｎ　Ｄａｔａ）等も含む。

[0024] アミノ酸配列情報１８は、抗体１６を構成するアミノ酸残基のペプチド結合の順番を、アミノ酸残基を表す１文字のアルファベットの略号を用いて、アミノ末端からカルボキシル末端に向かって記述したものである。抗体１６を構成するアミノ酸残基は大体４５０個程度あるので、アミノ酸配列情報１８のアルファベットも４５０個程度の連なりとなる。略号は、例えば「Ｅ」はグルタミン酸、「Ｌ」はロイシン、「Ｇ」はグリシンである。こうしたアミノ酸残基の配列は一次構造とも呼ばれる。

[0025] 予測要求１５を受信した場合、情報処理サーバ１０は、アミノ酸配列情報１８を元に性状情報１７の予測を行う。性状情報１７は、ここでは抗体１６の品質が最も安定的に維持される抗体１６の保存溶液の水素イオン指数、すなわち抗体１６の至適な水素イオン指数である。情報処理サーバ１０は、予測要求１５の送信元のユーザ端末１１に性状情報１７を配信する。性状情報１７を受信した場合、ユーザ端末１１は、性状情報１７をユーザ端末１１のディスプレイ２９Ｂ（図１３参照）に表示し、性状情報１７をユーザＵの閲覧に供する。

[0026] 一例として図２に示すように、抗体１６は、基本的には、４本のポリペプチド鎖、すなわち２本の同一の重鎖ＨＣ（Ｈｅａｖｙ　Ｃｈａｉｎ）と、２本の同一の軽鎖ＬＣ（Ｌｉｇｈｔ　Ｃｈａｉｎ）とを有する。抗体１６は、これら２本の重鎖ＨＣと２本の軽鎖ＬＣとがジスルフィド結合ＤＢ（Ｄｉｓ

ｕｌｆｉｄｅ　Ｂｏｎｄ）により結合された構成であり、左右対称なＹ字形状である。

[0027] 重鎖ＨＣは、可変ドメインＶＨ（Ｖａｒｉａｂｌｅ　ｄｏｍａｉｎ，Ｈｅａｖｙ　Ｃｈａｉｎ）と、定常ドメインＣＨ（Ｃｏｎｓｔａｎｔ　ｄｏｍａｉｎ，Ｈｅａｖｙ　Ｃｈａｉｎ）１、ＣＨ２、およびＣＨ３とで構成される。軽鎖ＬＣは、可変ドメインＶＬ（Ｖａｒｉａｂｌｅ　ｄｏｍａｉｎ，Ｌｉｇｈｔ　Ｃｈａｉｎ）と定常ドメインＣＬ（Ｃｏｎｓｔａｎｔ　ｄｏｍａｉｎ，Ｌｉｇｈｔ　Ｃｈａｉｎ）とで構成される。重鎖ＨＣの定常ドメインＣＨ１およびＣＨ２は、ヒンジ領域ＨＲ（Ｈｉｎｇｅ　Ｒｅｇｉｏｎ）によって繋がれている。可変ドメインＶＨおよびＶＬは、合わせて可変領域と呼ばれる。定常ドメインＣＨ１～ＣＨ３、およびＣＬは、合わせて定常領域と呼ばれる。可変ドメインＶＨおよびＶＬ、定常ドメインＣＨ１～ＣＨ３、およびＣＬは、本開示の技術に係る「区分領域」の一例である。

[0028] 可変ドメインＶＨには、３つの相補性決定領域ＣＤＲ－Ｈ（Ｃｏｍｐｌｅｍｅｎｔａｒｉｔｙ　Ｄｅｔｅｒｍｉｎｉｎｇ　Ｒｅｇｉｏｎ，Ｈｅａｖｙ　Ｃｈａｉｎ）１、ＣＤＲ－Ｈ２、およびＣＤＲ－Ｈ３が含まれている。また、可変ドメインＶＬにも、３つの相補性決定領域ＣＤＲ－Ｌ（Ｃｏｍｐｌｅｍｅｎｔａｒｉｔｙ　Ｄｅｔｅｒｍｉｎｉｎｇ　Ｒｅｇｉｏｎ，Ｌｉｇｈｔ　Ｃｈａｉｎ）１、ＣＤＲ－Ｌ２、およびＣＤＲ－Ｌ３が含まれている。これらの相補性決定領域ＣＤＲ－Ｈ１～ＣＤＲ－Ｈ３およびＣＤＲ－Ｌ１～ＣＤＲ－Ｌ３は抗原との結合部位であり、超可変領域とも呼ばれる。相補性決定領域ＣＤＲ－Ｈ１～ＣＤＲ－Ｈ３およびＣＤＲ－Ｌ１～ＣＤＲ－Ｌ３も、本開示の技術に係る「区分領域」の一例である。

[0029] 可変ドメインＶＨおよびＶＬと、定常ドメインＣＨ１およびＣＬとで構成される領域は、フラグメント抗原結合領域ＦａｂＲである。定常ドメインＣＨ２およびＣＨ３とヒンジ領域ＨＲの一部とで構成される領域は、フラグメント結晶化可能領域ＦｃＲである。可変ドメインＶＨおよびＶＬで構成される領域は、フラグメント可変領域ＦｖＲである。周知のように、フラグメン

ト抗原結合領域ＦａｂＲとフラグメント結晶化可能領域ＦｃＲは、パパイン消化により産出することが可能である。同様に、フラグメント可変領域ＦｖＲは、ペプシン消化により産出することが可能である。

[0030] 一例として図３に示すように、情報処理サーバ１０およびユーザ端末１１を構成するコンピュータは、基本的には同じ構成であり、ストレージ２５、メモリ２６、ＣＰＵ（Ｃｅｎｔｒａｌ Ｐｒｏｃｅｓｓｉｎｇ Ｕｎｉｔ）２７、通信部２８、ディスプレイ２９、および入力デバイス３０を備えている。これらはバスライン３１を介して相互接続されている。

[0031] ストレージ２５は、情報処理サーバ１０およびユーザ端末１１を構成するコンピュータに内蔵、またはケーブル、ネットワークを通じて接続されたハードディスクドライブである。もしくはストレージ２５は、ハードディスクドライブを複数台連装したディスクアレイである。ストレージ２５には、オペレーティングシステム等の制御プログラム、各種アプリケーションプログラム（以下、ＡＰ（Ａｐｐｌｉｃａｔｉｏｎ Ｐｒｏｇｒａｍ）と表記する）、およびこれらのプログラムに付随する各種データ等が記憶されている。なお、ハードディスクドライブに代えてソリッドステートドライブを用いてもよい。

[0032] メモリ２６は、ＣＰＵ２７が処理を実行するためのワークメモリである。ＣＰＵ２７は、ストレージ２５に記憶されたプログラムをメモリ２６へロードして、プログラムにしたがった処理を実行する。これによりＣＰＵ２７は、コンピュータの各部を統括的に制御する。ＣＰＵ２７は、本開示の技術に係る「プロセッサ」の一例である。なお、メモリ２６は、ＣＰＵ２７に内蔵されていてもよい。

[0033] 通信部２８は、外部装置との各種情報の伝送制御を行う。ディスプレイ２９は各種画面を表示する。各種画面にはＧＵＩ(Ｇｒａｐｈｉｃａｌ Ｕｓｅｒ Ｉｎｔｅｒｆａｃｅ)による操作機能が備えられる。情報処理サーバ１０およびユーザ端末１１を構成するコンピュータは、各種画面を通じて、入力デバイス３０からの操作指示の入力を受け付ける。入力デバイス３０は、キ

ーボード、マウス、タッチパネル、および音声入力用のマイク等である。

[0034] なお、以下の説明では、情報処理サーバ１０を構成するコンピュータの各部（ストレージ２５およびＣＰＵ２７）には添え字の「Ａ」を、ユーザ端末１１を構成するコンピュータの各部（ストレージ２５、ＣＰＵ２７、ディスプレイ２９、および入力デバイス３０）には添え字の「Ｂ」をそれぞれ符号に付して区別する。

[0035] 一例として図４に示すように、情報処理サーバ１０のストレージ２５Ａには、作動プログラム３５が記憶されている。作動プログラム３５は、コンピュータを情報処理サーバ１０として機能させるためのＡＰである。すなわち、作動プログラム３５は、本開示の技術に係る「情報処理装置の作動プログラム」の一例である。ストレージ２５Ａには、特徴量抽出モデル３６、設定補助情報３７、および性状情報予測モデル３８も記憶されている。性状情報予測モデル３８は、本開示の技術に係る「機械学習モデル」の一例である。

[0036] 作動プログラム３５が起動されると、情報処理サーバ１０を構成するコンピュータのＣＰＵ２７は、メモリ２６等と協働して、受付部４０、リードライト（以下、ＲＷ（Ｒｅａｄ　Ｗｒｉｔｅ）と略す）制御部４１、構造情報導出部４２、特徴量抽出部４３、重み設定部４４、前処理部４５、予測部４６、および画面配信制御部４７として機能する。

[0037] 受付部４０は、ユーザ端末１１からの各種要求を受け付ける。特に受付部４０は、ユーザ端末１１からの予測要求１５を受け付ける。前述のように、予測要求１５はアミノ酸配列情報１８を含んでいる。このため、受付部４０は、予測要求１５を受け付けることで、アミノ酸配列情報１８を取得していることになる。受付部４０は、アミノ酸配列情報１８をＲＷ制御部４１に出力する。また、受付部４０は、図示省略したユーザ端末１１の端末ＩＤを画面配信制御部４７に出力する。

[0038] ＲＷ制御部４１は、ストレージ２５Ａへの各種データの記憶、およびストレージ２５Ａ内の各種データの読み出しを制御する。例えばＲＷ制御部４１は、受付部４０からのアミノ酸配列情報１８をストレージ２５Ａに記憶する

。また、ＲＷ制御部４１は、アミノ酸配列情報１８をストレージ２５Ａから読み出し、読み出したアミノ酸配列情報１８を構造情報導出部４２および特徴量抽出部４３に出力する。

[0039] ＲＷ制御部４１は、特徴量抽出モデル３６をストレージ２５Ａから読み出し、読み出した特徴量抽出モデル３６を特徴量抽出部４３に出力する。また、ＲＷ制御部４１は、設定補助情報３７をストレージ２５Ａから読み出し、読み出した設定補助情報３７を重み設定部４４に出力する。さらに、ＲＷ制御部４１は、性状情報予測モデル３８をストレージ２５Ａから読み出し、読み出した性状情報予測モデル３８を予測部４６に出力する。

[0040] 構造情報導出部４２は、アミノ酸配列情報１８に基づいて、抗体１６に固有のアミノ酸配列の構造情報５０を導出する。構造情報導出部４２は、構造情報５０を重み設定部４４に出力する。

[0041] 特徴量抽出部４３は、アミノ酸配列情報１８を特徴量抽出モデル３６に適用することで、抗体１６のアミノ酸配列を構成する各アミノ酸残基の特徴量の集合である特徴量群５１を抽出する。特徴量抽出部４３は、特徴量群５１を重み設定部４４に出力する。

[0042] 重み設定部４４は、設定補助情報３７を参照しつつ、構造情報５０に基づいて、特徴量群５１を構成する各特徴量の重みを設定する。重み設定部４４は、設定した重みを特徴量群５１に登録して、特徴量群５１を重み付き特徴量群５１Ｗとする。重み設定部４４は、重み付き特徴量群５１Ｗを前処理部４５に出力する。

[0043] 前処理部４５は、重み付き特徴量群５１Ｗに登録された特徴量に対して重みに応じた前処理を施し、これにより集約特徴量５２を生成する。前処理部４５は、集約特徴量５２を予測部４６に出力する。

[0044] 予測部４６は、集約特徴量５２を性状情報予測モデル３８に適用することで、性状情報予測モデル３８に性状情報１７を予測させる。予測部４６は、性状情報１７を画面配信制御部４７に出力する。

[0045] 画面配信制御部４７は、各種画面をユーザ端末１１に配信する制御を行う

。具体的には、画面配信制御部４７は、各種画面を、例えばＸＭＬ（Ｅｘｔｅｎｓｉｂｌｅ　Ｍａｒｋｕｐ　Ｌａｎｇｕａｇｅ）等のマークアップ言語によって作成されるウェブ配信用の画面データの形式で、各種要求の送信元のユーザ端末１１に配信出力する。この際、画面配信制御部４７は、各種要求の送信元のユーザ端末１１を、受付部４０からの端末ＩＤを元に特定する。各種画面には、アミノ酸配列情報１８を入力するための情報入力画面７５（図１４参照）、および性状情報１７を表示するための予測結果表示画面８０（図１５参照）等がある。なお、ＸＭＬに代えて、ＪＳＯＮ（Ｊａｖａｓｃｒｉｐｔ（登録商標）　Ｏｂｊｅｃｔ　Ｎｏｔａｔｉｏｎ）等の他のデータ記述言語を利用してもよい。

[0046] 一例として図５に示すように、構造情報５０は、抗体１６のアミノ酸配列を構成する各アミノ酸残基について、アミノ酸残基の機能的な区分領域と、アミノ酸配列が形成する立体構造の３次元的な形状（以下、３次元形状と表記する）とが登録された情報である。すなわち、構造情報５０は、本開示の技術に係る「区分領域情報」および「３次元形状情報」の一例であり、さらには「立体構造情報」の一例である。

[0047] 区分領域には、図２で示した可変ドメインＶＨおよびＶＬ、定常ドメインＣＨ１～ＣＨ３、およびＣＬ、あるいは相補性決定領域ＣＤＲ－Ｈ１～ＣＤＲ－Ｈ３およびＣＤＲ－Ｌ１～ＣＤＲ－Ｌ３のいずれかが登録される。３次元形状には、直線形状およびループ形状のいずれかが登録される。

[0048] 構造情報導出部４２は、例えば、区分領域が既知の抗体１６の集合とのマルチプルアライメントによって、区分領域が未知の抗体１６の区分領域を特定する。なお、区分領域は、Ｘ線結晶構造解析等の周知の技術で抗体１６を解析することで得られる他、論文情報、あるいはＰｒｏｔｅｉｎ　Ｄａｔａ　Ｂａｎｋといったデータベースからも得ることができる。

[0049] また、構造情報導出部４２は、例えば、ＤｅｅｐＭｉｎｄ社が開発したタンパク質の立体構造の予測プログラムである「ＡｌｐｈａＦｏｌｄ２」にアミノ酸配列情報１８を入力する。そして、「ＡｌｐｈａＦｏｌｄ２」から各

アミノ酸残基の３次元座標を出力させ、当該３次元座標を元に３次元形状を特定する。

[0050] 一例として図６に示すように、特徴量抽出部４３は、アミノ酸配列情報１８を特徴量抽出モデル３６に入力し、これにより特徴量抽出モデル３６から特徴量群５１を出力させる。特徴量抽出モデル３６は、例えばニューラルネットワークのエンコーダ部、あるいはサポートベクターマシン（ＳＶＭ：Ｓｕｐｐｏｒｔ　Ｖｅｃｔｏｒ　Ｍａｃｈｉｎｅ）といった機械学習モデルである。特徴量抽出モデル３６の学習は、情報処理サーバ１０において行ってもよいし、情報処理サーバ１０以外の装置で行ってもよい。また、特徴量抽出モデル３６の学習は、運用後も継続して行ってもよい。

[0051] 特徴量には、特徴量Ａ、特徴量Ｂ、特徴量Ｃ、・・・、特徴量ＺＺＺと複数種類ある。抗体１６のアミノ酸配列を構成する各アミノ酸残基には、複数種類の特徴量が登録されている。例えばＮｏ．１のアミノ酸残基「Ｅ」には、特徴量ＡとしてＺＡ１、特徴量ＢとしてＺＢ１、特徴量ＣとしてＺＣ１、・・・、特徴量ＺＺＺとしてＺＺＺＺ１がそれぞれ登録されている。このように、各アミノ酸残基は、複数種類の特徴量で表すことができる。複数種類の特徴量は、いわば多次元のベクトルデータである。特徴量の種類（特徴量ベクトルの次元数）は、数百～数千である。なお、特徴量抽出モデル３６といった機械学習モデルに加えて、あるいは代えて、分子動力学（ＭＤ：Ｍｏｌｅｃｕｌａｒ　Ｄｙｎａｍｉｃｓ）法を用いて特徴量を抽出してもよい。

[0052] 一例として図７に示すように、設定補助情報３７は、区分領域と３次元形状の組み合わせ毎に、設定する重みが登録された情報である。区分領域は、相補性決定領域ＣＤＲか、相補性決定領域ＣＤＲ以外の領域かの２つのグループに分かれている。２つのグループは、３次元形状が直線形状かループ形状かでさらに分かれている。区分領域が相補性決定領域ＣＤＲで、３次元形状が直線形状の場合の重みは２、３次元形状がループ形状の場合の重みは２．５である。区分領域が相補性決定領域ＣＤＲ以外の領域で、３次元形状が直線形状の場合の重みは１、ループ形状の場合の重みは１．５である。区分

領域が相補性決定領域ＣＤＲである場合は、区分領域が相補性決定領域ＣＤＲ以外の領域である場合と比べて、重みが高く設定されている。また、３次元形状がループ形状である場合は、３次元形状が直線形状である場合と比べて、重みが高く設定されている。

[0053] 一例として図８に示すように、重み設定部４４は、設定補助情報３７を参照して、構造情報５０の区分領域と３次元形状の組み合わせに応じた重みを各アミノ酸残基に設定する。より詳しくは、重み設定部４４は、区分領域が相補性決定領域ＣＤＲであるアミノ酸残基の特徴量の重みを、区分領域が相補性決定領域ＣＤＲでないアミノ酸残基の特徴量の重みよりも高く設定する。また、重み設定部４４は、３次元形状がループ形状であるアミノ酸残基の特徴量の重みを、３次元形状が直線形状であるアミノ酸残基の特徴量の重みよりも高く設定する。例えばＮｏ．３のアミノ酸残基「Ｑ」は、区分領域が定常ドメインＣＨ１、３次元形状が直線形状であるので、重み設定部４４は重み１を設定する。また、例えばＮｏ．１００のアミノ酸残基「Ｓ」は、区分領域が相補性決定領域ＣＤＲ－Ｌ１、３次元形状がループ形状であるので、重み設定部４４は重み２．５を設定する。

[0054] 一例として図９に示すように、重み付き特徴量群５１Ｗは、特徴量群５１に重みの項を追加したものである。

[0055] 一例として図１０に示すように、前処理部４５は、特徴量を性状情報予測モデル３８に適用する前の前処理として、各特徴量の重みによる加重平均を算出し、集約特徴量５２とする。具体的には、前処理部４５は、特徴量ＡのＺＡ１、ＺＡ２、ＺＡ３、・・・の重みによる加重平均ＺＡ、特徴量ＢのＺＢ１、ＺＢ２、ＺＢ３、・・・の重みによる加重平均ＺＢ、特徴量ＣのＺＣ１、ＺＣ２、ＺＣ３、・・・の重みによる加重平均ＺＣ、・・・、特徴量ＺＺのＺＺＺＺ１、ＺＺＺＺ２、ＺＺＺＺ３、・・・の重みによる加重平均ＺＺＺＺをそれぞれ算出する。そして、加重平均ＺＡ、ＺＢ、ＺＣ、・・・、ＺＺＺＺの集合を集約特徴量５２として出力する。

[0056] 一例として図１１に示すように、予測部４６は、集約特徴量５２を性状情

報予測モデル３８に入力し、これにより性状情報予測モデル３８から性状情報１７を出力させる。

[0057] 一例として図１２に示すように、性状情報予測モデル３８はニューラルネットワーク６０により構築されている。ニューラルネットワーク６０は、周知のように入力層６１、中間層（隠れ層ともいう）６２、および出力層６３を有する。これら入力層６１、中間層６２、および出力層６３は、それぞれ複数のノードＮＤをもつ。入力層６１のノードＮＤと中間層６２のノードＮＤとの間、中間層６２内のノードＮＤの間、および中間層６２のノードＮＤと出力層６３のノードＮＤとの間には、各ノードＮＤの結合の強さを示す係数が設定される。出力層６３のノードＮＤには、線形関数、ＲｅＬｕ（Ｒｅｃｔｉｆｉｅｄ Ｌｉｎｅａｒ Ｕｎｉｔ）関数といった適当な活性化関数が設定されている。

[0058] 入力層６１の各ノードＮＤには、集約特徴量５２の各特徴量ＺＡ、ＺＢ、ＺＣ、・・・、ＺＺＺＺが入力される。また、出力層６３のノードＮＤからは、性状情報１７が出力される。なお、性状情報予測モデル３８は例示のニューラルネットワーク６０に限らず、決定木、勾配ブースティング決定木、ランダムフォレスト、サポートベクターマシン、ナイーブベイズといった他の機械学習モデルでもよい。

[0059] 性状情報予測モデル３８の学習は、情報処理サーバ１０において行ってもよいし、情報処理サーバ１０以外の装置で行ってもよい。また、性状情報予測モデル３８の学習は、運用後も継続して行ってもよい。

[0060] 一例として図１３に示すように、ユーザ端末１１のストレージ２５Ｂには、予測ＡＰ７０が記憶されている。予測ＡＰ７０は、ユーザＵによってユーザ端末１１にインストールされる。予測ＡＰ７０は、抗体１６の性状情報１７を予測するためのＡＰである。予測ＡＰ７０が起動された場合、ユーザ端末１１のＣＰＵ２７Ｂは、メモリ２６等と協働して、ブラウザ制御部７２として機能する。ブラウザ制御部７２は、予測ＡＰ７０の専用のウェブブラウザの動作を制御する。

[0061] ブラウザ制御部７２は、情報処理サーバ１０からの各種画面データに基づいて各種画面を再現し、再現した各種画面をディスプレイ２９Ｂに表示する。また、ブラウザ制御部７２は、各種画面を通じて、ユーザＵによって入力デバイス３０Ｂから入力される様々な操作指示を受け付ける。ブラウザ制御部７２は、予測要求１５をはじめとした、操作指示に応じた様々な要求を情報処理サーバ１０に送信する。

[0062] 予測ＡＰ７０が起動された場合、ブラウザ制御部７２の制御の下、一例として図１４に示す情報入力画面７５がディスプレイ２９Ｂに表示される。情報入力画面７５には、アミノ酸配列情報１８の入力ボックス７６が設けられている。入力ボックス７６には、アミノ酸配列情報１８を記述したり、アミノ酸配列情報１８のファイルをドロップしたりすることができる。

[0063] ユーザＵは、入力ボックス７６に所望のアミノ酸配列情報１８を入力した後、予測ボタン７７を選択する。予測ボタン７７が選択された場合、ブラウザ制御部７２は、入力ボックス７６に入力されたアミノ酸配列情報１８を含む予測要求１５を生成し、生成した予測要求１５を情報処理サーバ１０に送信する。

[0064] また、情報処理サーバ１０において性状情報１７の予測が行われた場合、ブラウザ制御部７２の制御の下、一例として図１５に示す予測結果表示画面８０がディスプレイ２９Ｂに表示される。予測結果表示画面８０には、性状情報１７、正確には性状情報１７を表すメッセージが表示される。このように、性状情報１７は、画面データの配信という形でユーザＵに提示される。

[0065] 予測結果表示画面８０の上部には、アミノ酸配列情報表示ボタン８１が設けられている。アミノ酸配列情報表示ボタン８１が選択された場合、アミノ酸配列情報１８の表示画面がポップアップ表示される。

[0066] また、予測結果表示画面８０の下部には、保存ボタン８２およびＯＫボタン８３が設けられている。保存ボタン８２が選択された場合、性状情報１７およびアミノ酸配列情報１８がユーザ端末１１のストレージ２５Ｂに関連付けて記憶される。ＯＫボタン８３が選択された場合、予測結果表示画面８０

の表示が消される。

[0067] 次に、上記構成による作用について、図１６のフローチャートを参照して説明する。まず、情報処理サーバ１０において作動プログラム３５が起動されると、図４で示したように、情報処理サーバ１０のＣＰＵ２７Ａは、受付部４０、ＲＷ制御部４１、構造情報導出部４２、特徴量抽出部４３、重み設定部４４、前処理部４５、予測部４６、および画面配信制御部４７として機能される。また、ユーザ端末１１において予測ＡＰ７０が起動されると、図１３で示したように、ユーザ端末１１のＣＰＵ２７Ｂは、ブラウザ制御部７２として機能される。

[0068] ユーザ端末１１のディスプレイ２９Ｂには、ブラウザ制御部７２の制御の下、図１４で示した情報入力画面７５が表示される。情報入力画面７５において、ユーザＵにより、入力ボックス７６に所望のアミノ酸配列情報１８が入力され、予測ボタン７７が選択された場合、ブラウザ制御部７２から情報処理サーバ１０に向けて予測要求１５が送信される。図１で示したように、予測要求１５には、アミノ酸配列情報１８と、ユーザ端末１１の端末ＩＤ等が含まれている。

[0069] 情報処理サーバ１０においては、予測要求１５が受付部４０にて受け付けられる（ステップＳＴ１００でＹＥＳ）。予測要求１５のアミノ酸配列情報１８は、受付部４０からＲＷ制御部４１に出力され、ＲＷ制御部４１の制御の下、ストレージ２５Ａに記憶される（ステップＳＴ１１０）。

[0070] ＲＷ制御部４１により、ストレージ２５Ａからアミノ酸配列情報１８が読み出される（ステップＳＴ１２０）。アミノ酸配列情報１８は、ＲＷ制御部４１から構造情報導出部４２および特徴量抽出部４３に出力される。

[0071] 構造情報導出部４２において、アミノ酸配列情報１８に基づいて、図５で示した構造情報５０が導出される（ステップＳＴ１３０）。構造情報５０は、構造情報導出部４２から重み設定部４４に出力される。

[0072] また、図６で示したように、特徴量抽出部４３において、アミノ酸配列情報１８が特徴量抽出モデル３６に入力され、これにより特徴量抽出モデル３

６から特徴量群５１が出力される（ステップＳＴ１４０）。特徴量群５１は、特徴量抽出部４３から重み設定部４４に出力される。

[0073] 次に、図８で示したように、重み設定部４４において、設定補助情報３７が参照されつつ、構造情報５０に基づいて特徴量の重みが設定される（ステップＳＴ１５０）。重み設定部４４では、設定された重みの項を含む重み付き特徴量群５１Ｗが生成される。重み付き特徴量群５１Ｗは、重み設定部４４から前処理部４５に出力される。

[0074] 図１０で示したように、前処理部４５において、前処理として、各特徴量の重みによる加重平均が算出され、これにより集約特徴量５２が生成される（ステップＳＴ１６０）。集約特徴量５２は、前処理部４５から予測部４６に出力される。

[0075] 図１１で示したように、予測部４６において、集約特徴量５２が性状情報予測モデル３８に入力され、これにより性状情報予測モデル３８から性状情報１７が出力される（ステップＳＴ１７０）。性状情報１７は、予測部４６から画面配信制御部４７に出力される。

[0076] 画面配信制御部４７により、性状情報１７を表すメッセージを含む予測結果表示画面８０の画面データが生成される。予測結果表示画面８０の画面データは、画面配信制御部４７の制御の下、予測要求１５の送信元のユーザ端末１１に配信される（ステップＳＴ１８０）。

[0077] ユーザ端末１１のディスプレイ２９Ｂには、図１５で示した予測結果表示画面８０が表示される。ユーザＵは、予測結果表示画面８０を閲覧し、性状情報１７の至適な水素イオン指数を確認する。

[0078] 以上説明したように、情報処理サーバ１０のＣＰＵ２７Ａは、構造情報導出部４２、特徴量抽出部４３、重み設定部４４、および予測部４６を備える。構造情報導出部４２は、抗体１６に固有のアミノ酸配列の構造情報５０を導出することで取得する。特徴量抽出部４３は、抗体１６のアミノ酸配列を構成するアミノ酸残基の特徴量を抽出することで取得する。重み設定部４４は、構造情報５０に基づいて特徴量の重みを設定する。予測部４６は、重み

が設定された特徴量を性状情報予測モデル３８に適用することで、性状情報予測モデル３８に抗体１６の性状を示す性状情報１７を予測させる。

[0079] 抗体１６の構造情報５０に基づいた特徴量の重みを設定することで、抗体１６の構造でその性状に深く関わると思われる重要なアミノ酸残基の特徴量を、性状情報１７の予測により役立てることができる。したがって、抗体１６の性状を示す性状情報１７を高い精度で予測することが可能となる。

[0080] 構造情報５０は、抗体１６が変異体であるか否かに関わらず導出可能な情報である。このため本開示の技術は、特開２０２１－１５８９７３号公報に記載の技術のように変異体のみに限定的に用いられるものではなく、変異体はもちろんのこと、変異体ではない一般的な抗体１６にも汎用的に用いることができる。

[0081] 図５で示したように、構造情報５０は、アミノ酸残基の機能的な区分領域を示す区分領域情報である。そして、タンパク質は抗体１６であり、区分領域は相補性決定領域ＣＤＲを含む。図７および図８で示したように、重み設定部４４は、区分領域が相補性決定領域ＣＤＲであるアミノ酸残基の特徴量の重みを、区分領域が相補性決定領域ＣＤＲでないアミノ酸残基の特徴量の重みよりも高く設定する。

[0082] 相補性決定領域ＣＤＲは、前述のように抗原との結合部位であり、抗体１６の機能を決定付ける領域である。また、相補性決定領域ＣＤＲは、抗体１６の保存溶液といった溶媒に一般的に多く露出する。このため、構造情報５０を、相補性決定領域ＣＤＲを含む区分領域を示す区分領域情報とし、区分領域が相補性決定領域ＣＤＲであるアミノ酸残基の特徴量の重みを、区分領域が相補性決定領域ＣＤＲでないアミノ酸残基の特徴量の重みよりも高く設定すれば、性状情報１７をさらに高い精度で予測することが可能となる。性状情報１７が、抗体１６の至適な水素イオン指数といった抗体１６の保存溶液に関わるものであった場合は特に、その効果は高まる。

[0083] また、図５で示したように、構造情報５０は、アミノ酸配列の立体構造情報であり、立体構造情報は、アミノ酸配列の３次元形状を示す３次元形状情

報である。そして、３次元形状は、ループ形状および直線形状を含む。図７および図８で示したように、重み設定部４４は、３次元形状がループ形状であるアミノ酸残基の特徴量の重みを、３次元形状が直線形状であるアミノ酸残基の特徴量の重みよりも高く設定する。

[0084] ３次元形状がループ形状であった場合、３次元形状が直線形状であった場合よりも、そのアミノ酸残基は溶媒に一般的に多く露出する。このため、構造情報５０を、ループ形状および直線形状を含むアミノ酸配列の３次元形状を示す３次元形状情報とし、３次元形状がループ形状であるアミノ酸残基の特徴量の重みを、３次元形状が直線形状であるアミノ酸残基の特徴量の重みよりも高く設定すれば、性状情報１７をさらに高い精度で予測することが可能となる。性状情報１７が、抗体１６の至適な水素イオン指数といった抗体１６の保存溶液に関わるものであった場合は特に、その効果は高まる。

[0085] 前処理部４５は、アミノ酸残基の特徴量に対して、重みに応じた前処理を施す。具体的には図１０で示したように、前処理は、アミノ酸残基の特徴量の重みによる加重平均を算出する処理である。このため、重みを反映した特徴量を性状情報予測モデル３８に適用することができる。

[0086] 至適な水素イオン指数は、抗体１６の品質を安定的に維持するためには非常に重要である。このため、性状情報１７を抗体１６の至適な水素イオン指数とすれば、抗体１６の品質の安定的な維持に寄与することができ、結果的に抗体１６を有効成分とするバイオ医薬品の薬効を十分に発揮させることができる。

[0087] タンパク質として抗体１６を含むバイオ医薬品は、抗体医薬品と呼ばれ、癌、糖尿病、関節リウマチといった慢性疾患の治療をはじめとして、血友病、クローン病といった希少疾患の治療にも幅広く用いられている。このため、タンパク質を抗体１６とした本例によれば、色々な疾患の治療に幅広く用いられている抗体医薬品の安定的な供給に寄与することができる。

[0088] 立体構造情報は、例示の３次元形状情報に限らない。一例として図１７に示す構造情報９０のように、各アミノ酸残基の３次元座標、および抗体１６

の重心座標の情報であってもよい。各アミノ酸残基の３次元座標、および抗体１６の重心座標は、「ＡｌｐｈａＦｏｌｄ２」にアミノ酸配列情報１８を入力して得られる。

[0089] この場合、重み設定部４４は、一例として図１８に示す設定補助情報９２を参照して重みを設定する。設定補助情報９２は、設定補助情報３７の３次元形状の項に代えて、重心からの距離の項を有する。重心からの距離は、各アミノ酸残基の３次元座標と抗体１６の重心座標とから算出することができる。区分領域は、設定補助情報３７の場合と同じく、相補性決定領域ＣＤＲか、相補性決定領域ＣＤＲ以外の領域かの２つのグループに分かれている。２つのグループは、重心からの距離が閾値未満か閾値以上かでさらに分かれている。区分領域が相補性決定領域ＣＤＲで、重心からの距離が閾値未満の場合の重みは２、重心からの距離が閾値以上の場合の重みは２．５である。区分領域が相補性決定領域ＣＤＲ以外の領域で、重心からの距離が閾値未満の場合の重みは１、閾値以上の場合の重みは１．５である。重心からの距離が閾値以上である場合は、重心からの距離が閾値未満である場合と比べて、重みが高く設定されている。

[0090] 重心からの距離が離れている程、そのアミノ酸残基は溶媒に一般的に多く露出する。このため、立体構造情報を、各アミノ酸残基の３次元座標と抗体１６の重心座標の情報とし、重心からの距離が閾値以上であるアミノ酸残基の特徴量の重みを、重心からの距離が閾値未満であるアミノ酸残基の特徴量の重みよりも高く設定すれば、性状情報１７をさらに高い精度で予測することが可能となる。性状情報１７が、抗体１６の至適な水素イオン指数といった抗体１６の保存溶液に関わるものであった場合は特に、その効果は高まる。

[0091] また、立体構造情報は、一例として図１９に示す構造情報９５のように、各アミノ酸残基の溶媒露出表面積（ＳＡＳＡ：Ｓｏｌｖｅｎｔ Ａｃｃｅｓｓｉｂｌｅ Ｓｕｒｆａｃｅ Ａｒｅａ）の情報であってもよい。溶媒露出表面積は、各アミノ酸残基の３次元座標に基づいて導出される。

[0092] この場合、重み設定部４４は、一例として図２０に示す設定補助情報９７を参照して重みを設定する。設定補助情報９７は、溶媒露出表面積が閾値未満の場合の重みを０、閾値以上の場合の重みを１としたものである。重みが０に設定されたアミノ酸残基の特徴量は、前処理の加重平均の算出の際には加算されない。この例からも分かるように、重み設定部４４が設定する重みは、０を含んでいてもよい。

[0093] 溶媒露出表面積が大きい程、名の通り、そのアミノ酸残基は溶媒に多く露出する。このため、立体構造情報を溶媒露出表面積の情報とし、溶媒露出表面積が閾値以上であるアミノ酸残基の特徴量の重みを、溶媒露出表面積が閾値未満であるアミノ酸残基の特徴量の重みよりも高く設定すれば、性状情報１７をさらに高い精度で予測することが可能となる。性状情報１７が、抗体１６の至適な水素イオン指数といった抗体１６の保存溶液に関わるものであった場合は特に、その効果は高まる。

[0094] ［第２実施形態］

一例として図２１に示すように、第２実施形態の前処理部１００は、学習前の性状情報予測モデル１０１に対して重みに応じた前処理を施し、性状情報予測モデル１０１を前処理済み性状情報予測モデル１０１Ａとする。上記第１実施形態の性状情報予測モデル３８は、入力層６１の各ノードＮＤに集約特徴量５２の各特徴量が入力されるものであったが、本第２実施形態の性状情報予測モデル１０１は、入力層６１の各ノードＮＤに特徴量群５１の各特徴量が入力される。

[0095] 前処理部１００は、前処理として、重みに応じた初期係数を学習前の性状情報予測モデル１０１の各ノードＮＤ間に設定する。具体的には、重みが相対的に高い特徴量がより重要に扱われるように、重みが相対的に高い特徴量に関わるノードＮＤ間の初期係数を、重みが相対的に低い特徴量に関わるノードＮＤ間の係数よりも高く設定する。こうして初期係数が設定された前処理済み性状情報予測モデル１０１Ａを学習することで、予測部４６において用いる性状情報予測モデル１０１Ｘを得る。

[0096] このように、第２実施形態においては、前処理は、重みに応じた初期係数を性状情報予測モデル１０１に設定する処理である。このため、少ない教師データで前処理済み性状情報予測モデル１０１Ａの学習を収束させることができ、比較的高い予測精度の性状情報予測モデル１０１Ｘを容易に得ることができる。

[0097] 情報処理サーバ１０とは別の装置で構造情報を導出し、ネットワーク１２等を介して、別の装置から情報処理サーバ１０に構造情報を送信する態様としてもよい。同様に、特徴量も情報処理サーバ１０とは別の装置で抽出し、ネットワーク１２等を介して、別の装置から情報処理サーバ１０に特徴量を送信する態様としてもよい。

[0098] アミノ酸残基の特徴量としては、空間凝集傾向（ＳＡＰ：Ｓｐａｔｉａｌ Ａｇｇｒｅｇａｔｉｏｎ Ｐｒｏｐｅｎｓｉｔｙ）、および空間電荷マップ（ＳＣＭ：Ｓｐａｔｉａｌ Ｃｈａｒｇｅ Ｍａｐ）等を採用してもよい。

[0099] 性状情報１７は例示の至適な水素イオン指数に限らない。抗体１６の品質が最も安定的に維持される抗体１６の保存溶液の粘度、Ｌ－アルギニン塩酸塩、精製白糖といった保存溶液に添加する添加剤の種類、あるいは添加剤の濃度等でもよい。また、抗体１６の凝集のしやすさを表す指標、抗体１６の疎水性を表す指標、抗体１６の毒性の有無等でもよい。

[0100] タンパク質は例示の抗体１６に限らない。ペプチド、核酸等でもよい。また、サイトカイン（インターフェロン、インターロイキン等）、あるいはホルモン（インスリン、グルカゴン、卵胞刺激ホルモン、エリスロポエチン等）、成長因子（ＩＧＦ（Ｉｎｓｕｌｉｎ－Ｌｉｋｅ Ｇｒｏｗｔｈ Ｆａｃｔｏｒ）－１、ｂＦＧＦ（Ｂａｓｉｃ Ｆｉｂｒｏｂｌａｓｔ Ｇｒｏｗｔｈ Ｆａｃｔｏｒ）等）、血液凝固因子（第７因子、第８因子、第９因子等）、酵素（リソソーム酵素、ＤＮＡ（Ｄｅｏｘｙｒｉｂｏｎｕｃｌｅｉｃ Ａｃｉｄ）分解酵素等）、Ｆｃ（Ｆｒａｇｍｅｎｔ Ｃｒｙｓｔａｌｌｉｚａｂｌｅ）融合タンパク質、受容体、アルブミン、タンパク質ワクチンでも

よい。また、抗体としては、バイスペシフィック抗体、抗体薬物複合体（Ａｎｔｉｂｏｄｙ－ｄｒｕｇ　ｃｏｎｊｕｇａｔｅ）、低分子抗体、糖鎖改変抗体等も含まれる。

[0101] 性状情報１７を表すメッセージを含む予測結果表示画面８０の画面データをユーザ端末１１に配信するのではなく、性状情報１７自体を情報処理サーバ１０からユーザ端末１１に配信してもよい。

[0102] 性状情報１７をユーザＵの閲覧に供する態様は、例示の予測結果表示画面８０に限らない。性状情報１７の印刷物をユーザＵに提供してもよいし、性状情報１７を添付した電子メールをユーザＵの携帯端末に送信してもよい。

[0103] 情報処理サーバ１０は、各製薬施設に設置されていてもよいし、製薬施設からは独立したデータセンターに設置されていてもよい。また、ユーザ端末１１が情報処理サーバ１０の各処理部４０～４７の一部または全ての機能を担ってもよい。

[0104] 本開示の技術に係る情報処理サーバ１０を構成するコンピュータのハードウェア構成は種々の変形が可能である。例えば、情報処理サーバ１０を、処理能力および信頼性の向上を目的として、ハードウェアとして分離された複数台のコンピュータで構成することも可能である。例えば、受付部４０、ＲＷ制御部４１、構造情報導出部４２、および特徴量抽出部４３の機能と、重み設定部４４、前処理部４５、予測部４６、および画面配信制御部４７の機能とを、２台のコンピュータに分散して担わせる。この場合は２台のコンピュータで情報処理サーバ１０を構成する。

[0105] このように、情報処理サーバ１０のコンピュータのハードウェア構成は、処理能力、安全性、および信頼性等の要求される性能に応じて適宜変更することができる。さらに、ハードウェアに限らず、作動プログラム３５および予測ＡＰ７０等のＡＰについても、安全性および信頼性の確保を目的として、二重化したり、あるいは、複数のストレージに分散して格納したりすることももちろん可能である。

[0106] 上記各実施形態において、例えば、受付部４０、ＲＷ制御部４１、構造情

報導出部４２、特徴量抽出部４３、重み設定部４４、前処理部４５および１００、予測部４６、画面配信制御部４７、並びにブラウザ制御部７２といった各種の処理を実行する処理部（Ｐｒｏｃｅｓｓｉｎｇ　Ｕｎｉｔ）のハードウェア的な構造としては、次に示す各種のプロセッサ（Ｐｒｏｃｅｓｓｏｒ）を用いることができる。各種のプロセッサには、上述したように、ソフトウェア（作動プログラム３５および予測ＡＰ７０）を実行して各種の処理部として機能する汎用的なプロセッサであるＣＰＵ２７Ａおよび２７Ｂに加えて、ＦＰＧＡ（Ｆｉｅｌｄ　Ｐｒｏｇｒａｍｍａｂｌｅ　Ｇａｔｅ　Ａｒｒａｙ）等の製造後に回路構成を変更可能なプロセッサであるプログラマブルロジックデバイス（Ｐｒｏｇｒａｍｍａｂｌｅ　Ｌｏｇｉｃ　Ｄｅｖｉｃｅ:ＰＬＤ）、ＡＳＩＣ（Ａｐｐｌｉｃａｔｉｏｎ　Ｓｐｅｃｉｆｉｃ　Ｉｎｔｅｇｒａｔｅｄ　Ｃｉｒｃｕｉｔ）等の特定の処理を実行させるために専用に設計された回路構成を有するプロセッサである専用電気回路等が含まれる。

[0107] １つの処理部は、これらの各種のプロセッサのうちの１つで構成されてもよいし、同種または異種の２つ以上のプロセッサの組み合わせ（例えば、複数のＦＰＧＡの組み合わせ、および／または、ＣＰＵとＦＰＧＡとの組み合わせ）で構成されてもよい。また、複数の処理部を１つのプロセッサで構成してもよい。

[0108] 複数の処理部を１つのプロセッサで構成する例としては、第１に、クライアントおよびサーバ等のコンピュータに代表されるように、１つ以上のＣＰＵとソフトウェアの組み合わせで１つのプロセッサを構成し、このプロセッサが複数の処理部として機能する形態がある。第２に、システムオンチップ（Ｓｙｓｔｅｍ　Ｏｎ　Ｃｈｉｐ:ＳｏＣ）等に代表されるように、複数の処理部を含むシステム全体の機能を１つのＩＣ（Ｉｎｔｅｇｒａｔｅｄ　Ｃｉｒｃｕｉｔ）チップで実現するプロセッサを使用する形態がある。このように、各種の処理部は、ハードウェア的な構造として、上記各種のプロセッサの１つ以上を用いて構成される。

[0109] さらに、これらの各種のプロセッサのハードウェア的な構造としては、より具体的には、半導体素子等の回路素子を組み合わせた電気回路（ｃｉｒｃｕｉｔｒｙ）を用いることができる。

[0110] 以上の記載から、下記の付記項に記載の技術を把握することができる。

[0111] ［付記項１］

プロセッサを備え、

前記プロセッサは、

タンパク質に固有のアミノ酸配列の構造情報と、前記アミノ酸配列を構成するアミノ酸残基の特徴量とを取得し、

前記構造情報に基づいて前記特徴量の重みを設定し、

前記重みが設定された前記特徴量を機械学習モデルに適用することで、前記機械学習モデルに前記タンパク質の性状を示す性状情報を予測させる、

情報処理装置。

［付記項２］

前記構造情報は、前記アミノ酸残基の機能的な区分領域を示す区分領域情報である付記項１に記載の情報処理装置。

［付記項３］

前記タンパク質は抗体であり、

前記区分領域は相補性決定領域を含み、

前記プロセッサは、

前記区分領域情報で示される前記区分領域が前記相補性決定領域である前記アミノ酸残基の特徴量の重みを、前記区分領域情報で示される前記区分領域が前記相補性決定領域でない前記アミノ酸残基の特徴量の重みよりも高く設定する付記項２に記載の情報処理装置。

［付記項４］

前記構造情報は、前記アミノ酸配列の立体構造情報である付記項１から付記項３のいずれか１項に記載の情報処理装置。

［付記項５］

前記立体構造情報は、前記アミノ酸配列の３次元的な形状を示す３次元形状情報である付記項４に記載の情報処理装置。

［付記項６］

前記３次元的な形状は、ループ形状および直線形状を含み、

前記プロセッサは、

前記３次元形状情報で示される前記３次元的な形状が前記ループ形状である前記アミノ酸残基の特徴量の重みを、前記３次元形状情報で示される前記３次元的な形状が前記直線形状である前記アミノ酸残基の特徴量の重みよりも高く設定する付記項５に記載の情報処理装置。

［付記項７］

前記プロセッサは、

前記アミノ酸残基の特徴量、または前記機械学習モデルに対して、前記重みに応じた前処理を施す付記項１から付記項６のいずれか１項に記載の情報処理装置。

［付記項８］

前記前処理は、前記アミノ酸残基の特徴量の前記重みによる加重平均を算出する処理である付記項７に記載の情報処理装置。

［付記項９］

前記前処理は、前記重みに応じた初期係数を前記機械学習モデルに設定する処理である付記項７に記載の情報処理装置。

［付記項１０］

前記性状情報は、前記タンパク質の至適な水素イオン指数である付記項１から付記項９のいずれか１項に記載の情報処理装置。

［付記項１１］

前記タンパク質は抗体である付記項１から付記項１０のいずれか１項に記載の情報処理装置。

[0112] 本開示の技術は、上述の種々の実施形態および／または種々の変形例を適宜組み合わせることも可能である。また、上記各実施形態に限らず、要旨を

逸脱しない限り種々の構成を採用し得ることはもちろんである。さらに、本開示の技術は、プログラムに加えて、プログラムを非一時的に記憶する記憶媒体、並びにプログラムを含むコンピュータプログラム製品にもおよぶ。

[0113] 以上に示した記載内容および図示内容は、本開示の技術に係る部分についての詳細な説明であり、本開示の技術の一例に過ぎない。例えば、上記の構成、機能、作用、および効果に関する説明は、本開示の技術に係る部分の構成、機能、作用、および効果の一例に関する説明である。よって、本開示の技術の主旨を逸脱しない範囲内において、以上に示した記載内容および図示内容に対して、不要な部分を削除したり、新たな要素を追加したり、置き換えたりしてもよいことはいうまでもない。また、錯綜を回避し、本開示の技術に係る部分の理解を容易にするために、以上に示した記載内容および図示内容では、本開示の技術の実施を可能にする上で特に説明を要しない技術常識等に関する説明は省略されている。

[0114] 本明細書において、「Ａおよび／またはＢ」は、「ＡおよびＢのうちの少なくとも１つ」と同義である。つまり、「Ａおよび／またはＢ」は、Ａだけであってもよいし、Ｂだけ

であってもよいし、ＡおよびＢの組み合わせであってもよい、という意味である。また、本明細書において、３つ以上の事柄を「および／または」で結び付けて表現する場合も、「Ａおよび／またはＢ」と同様の考え方が適用される。

[0115] 本明細書に記載された全ての文献、特許出願および技術規格は、個々の文献、特許出願および技術規格が参照により取り込まれることが具体的かつ個々に記された場合と同程度に、本明細書中に参照により取り込まれる。

請求の範囲

[請求項1] プロセッサを備え、

前記プロセッサは、

タンパク質に固有のアミノ酸配列の構造情報と、前記アミノ酸配列を構成するアミノ酸残基の特徴量とを取得し、

前記構造情報に基づいて前記特徴量の重みを設定し、

前記重みが設定された前記特徴量を機械学習モデルに適用することで、前記機械学習モデルに前記タンパク質の性状を示す性状情報を予測させる、

情報処理装置。

[請求項2] 前記構造情報は、前記アミノ酸残基の機能的な区分領域を示す区分領域情報である請求項１に記載の情報処理装置。

[請求項3] 前記タンパク質は抗体であり、

前記区分領域は相補性決定領域を含み、

前記プロセッサは、

前記区分領域情報で示される前記区分領域が前記相補性決定領域である前記アミノ酸残基の特徴量の重みを、前記区分領域情報で示される前記区分領域が前記相補性決定領域でない前記アミノ酸残基の特徴量の重みよりも高く設定する請求項２に記載の情報処理装置。

[請求項4] 前記構造情報は、前記アミノ酸配列の立体構造情報である請求項１に記載の情報処理装置。

[請求項5] 前記立体構造情報は、前記アミノ酸配列の３次元的な形状を示す３次元形状情報である請求項４に記載の情報処理装置。

[請求項6] 前記３次元的な形状は、ループ形状および直線形状を含み、

前記プロセッサは、

前記３次元形状情報で示される前記３次元的な形状が前記ループ形状である前記アミノ酸残基の特徴量の重みを、前記３次元形状情報で示される前記３次元的な形状が前記直線形状である前記アミノ酸残基

の特徴量の重みよりも高く設定する請求項５に記載の情報処理装置。

[請求項7] 前記プロセッサは、

前記アミノ酸残基の特徴量、または前記機械学習モデルに対して、前記重みに応じた前処理を施す請求項１に記載の情報処理装置。

[請求項8] 前記前処理は、前記アミノ酸残基の特徴量の前記重みによる加重平均を算出する処理である請求項７に記載の情報処理装置。

[請求項9] 前記前処理は、前記重みに応じた初期係数を前記機械学習モデルに設定する処理である請求項７に記載の情報処理装置。

[請求項10] 前記性状情報は、前記タンパク質の至適な水素イオン指数である請求項１に記載の情報処理装置。

[請求項11] 前記タンパク質は抗体である請求項１に記載の情報処理装置。

[請求項12] タンパク質に固有のアミノ酸配列の構造情報と、前記アミノ酸配列を構成するアミノ酸残基の特徴量とを取得すること、

前記構造情報に基づいて前記特徴量の重みを設定すること、並びに、

前記重みが設定された前記特徴量を機械学習モデルに適用することで、前記機械学習モデルに前記タンパク質の性状を示す性状情報を予測させること、

を含む情報処理装置の作動方法。

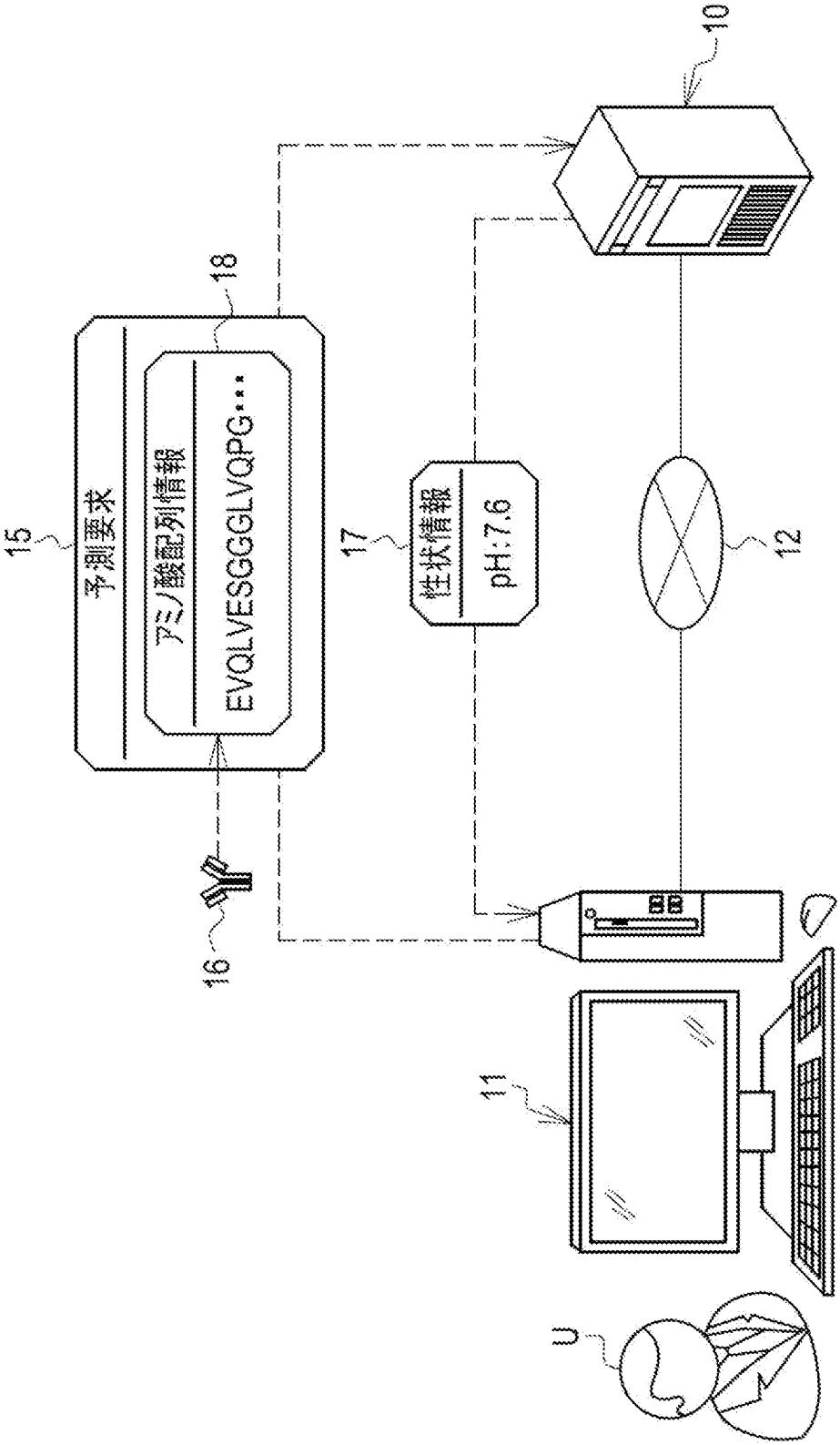
[請求項13] タンパク質に固有のアミノ酸配列の構造情報と、前記アミノ酸配列を構成するアミノ酸残基の特徴量とを取得すること、

前記構造情報に基づいて前記特徴量の重みを設定すること、並びに、

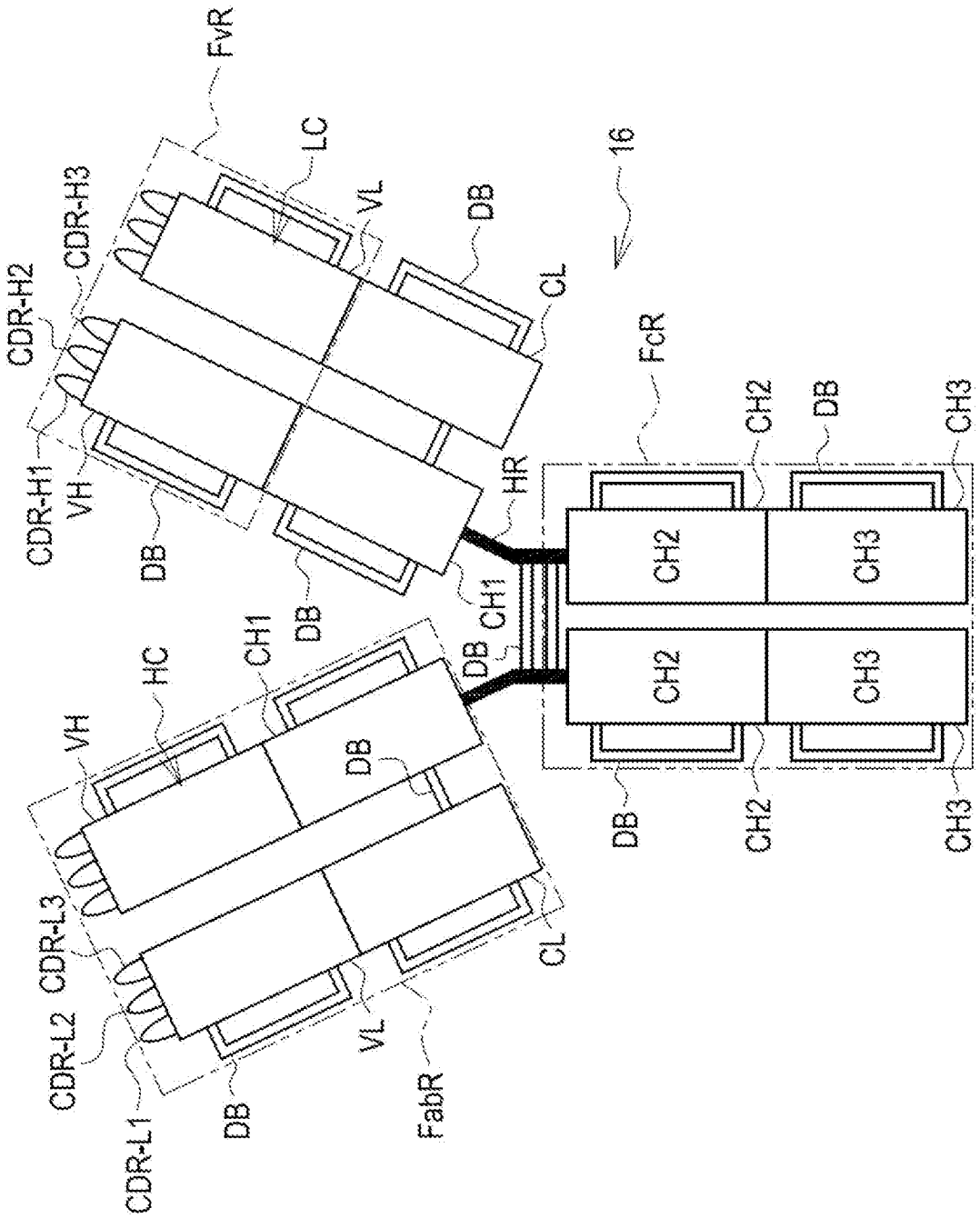
前記重みが設定された前記特徴量を機械学習モデルに適用することで、前記機械学習モデルに前記タンパク質の性状を示す性状情報を予測させること、

を含む処理をコンピュータに実行させる情報処理装置の作動プログラム。

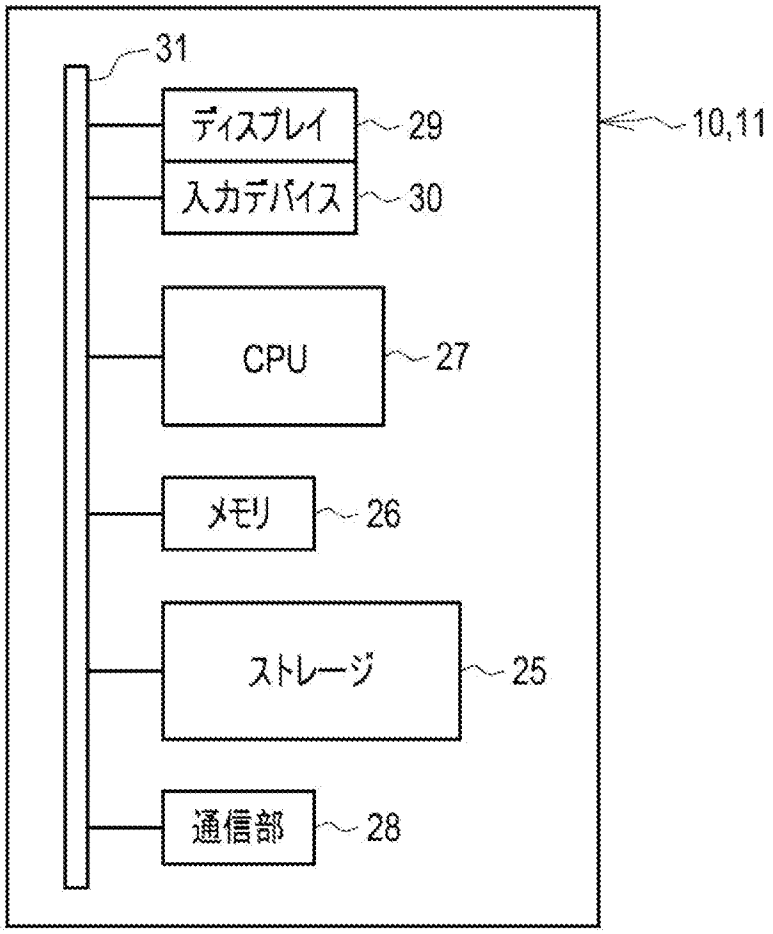
[図1]



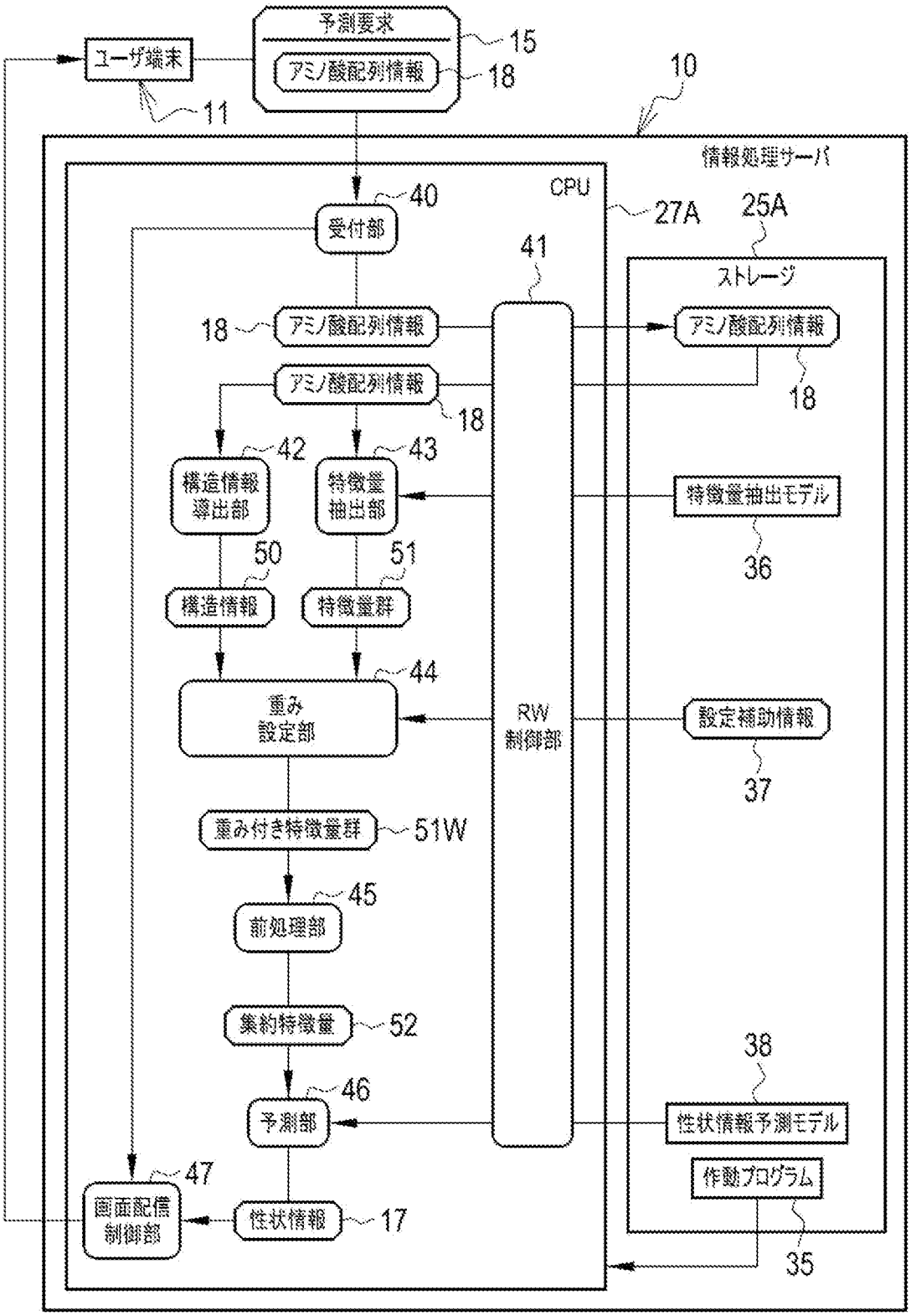
[図2]


FvR
LC
VL
DB
16
CDR-H3
CDR-H2
CL
FcR
CH2
DB
CH3
CDR-H1
VH
HR
DB
DB
CH1
CH2
CH3
DB
CH1
HC
VH
CH2
CH3
DB
DB
CH2
CH3
CDR-L3
CL
CDR-L2
VL
CDR-L1
DB
FabR

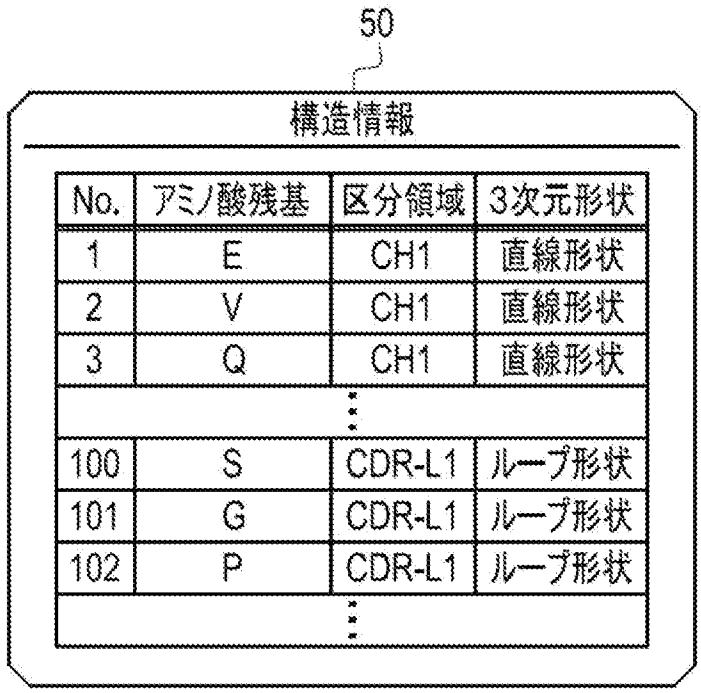
[図3]



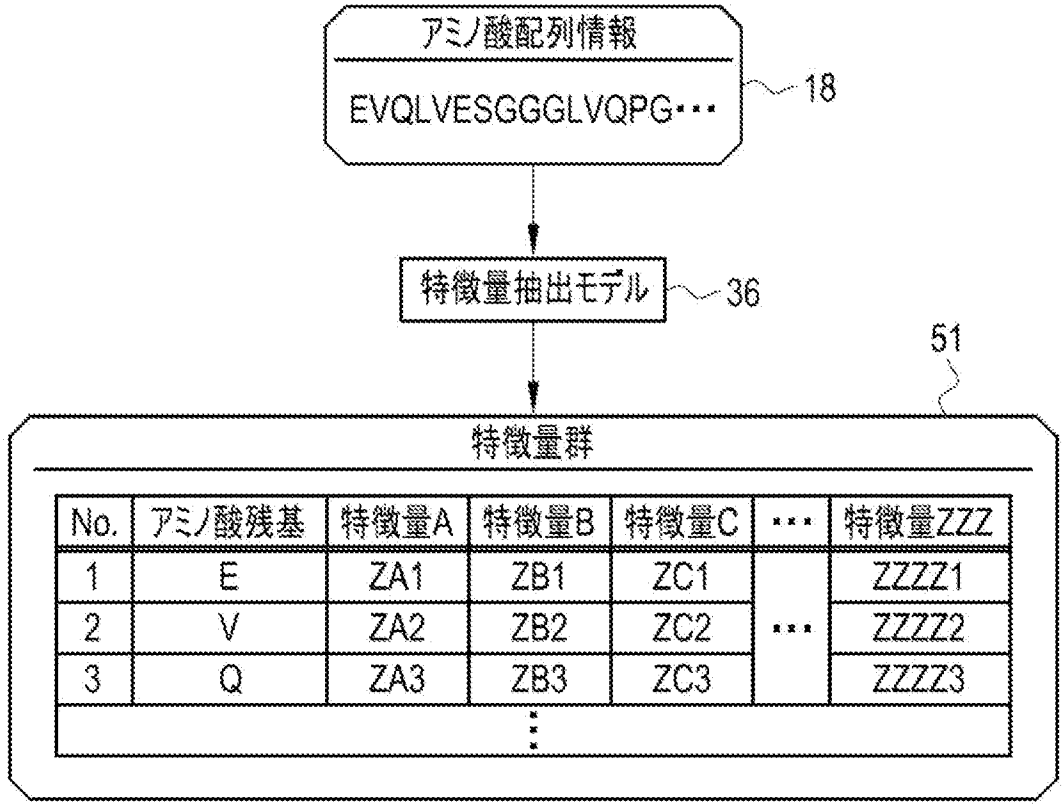
[図4]


予測要求
15
ユーザ端末
アミノ酸配列情報
18
11
10
情報処理サーバ
CPU
27A
25A
ストレージ
40
受付部
41
18
アミノ酸配列情報
アミノ酸配列情報
18
42
43
構造情報導出部
特徴量抽出部
特徴量抽出モデル
36
50
51
構造情報
特徴量群
44
重み設定部
RW制御部
設定補助情報
37
重み付き特徴量群
51W
45
前処理部
集約特徴量
52
38
46
予測部
性状情報予測モデル
作動プログラム
47
画面配信制御部
性状情報
17
35

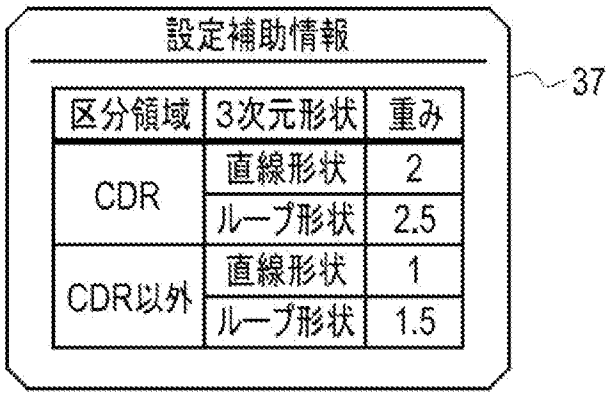
[図5]


50
構造情報
No. アミノ酸残基 区分領域 3次元形状
1 E CH1 直線形状
2 V CH1 直線形状
3 Q CH1 直線形状
100 S CDR-L1 ループ形状
101 G CDR-L1 ループ形状
102 P CDR-L1 ループ形状

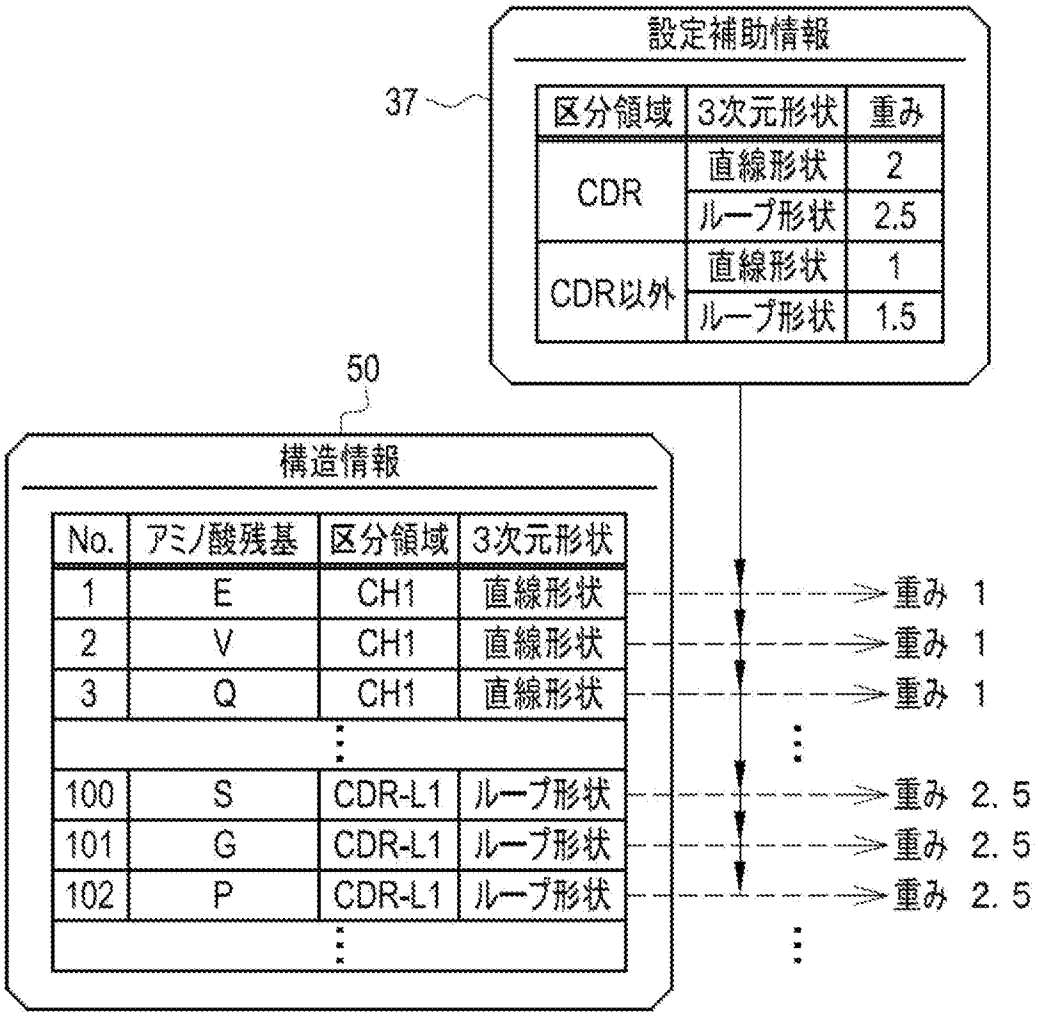
[図6]


アミノ酸配列情報
EVQLVESGGGLVQPG・・・
18
特徴量抽出モデル
36
51
特徴量群
No. アミノ酸残基 特徴量A 特徴量B 特徴量C ・・・ 特徴量ZZZ
1 E ZA1 ZB1 ZC1 ZZZZ1
2 V ZA2 ZB2 ZC2 ・・・ ZZZZ2
3 Q ZA3 ZB3 ZC3 ZZZZ3

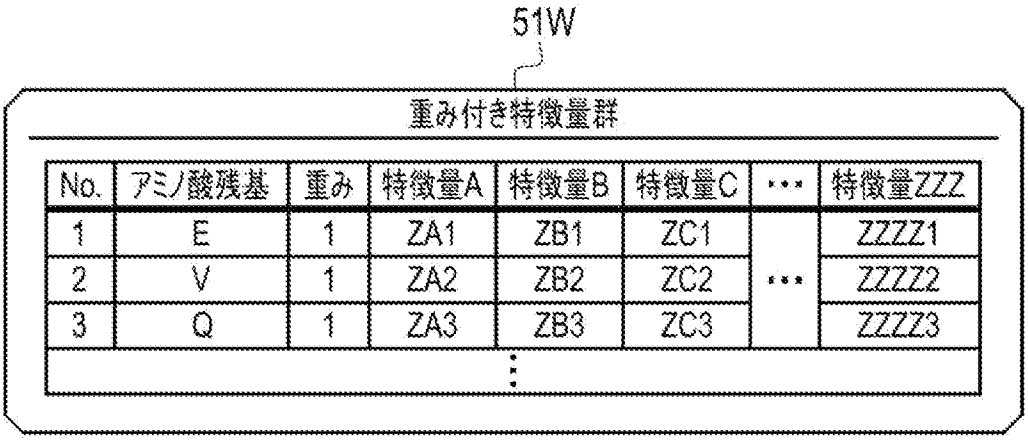
[図7]


設定補助情報
37
区分領域
3次元形状
重み
CDR
直線形状
2
ループ形状
2.5
CDR以外
直線形状
1
ループ形状
1.5

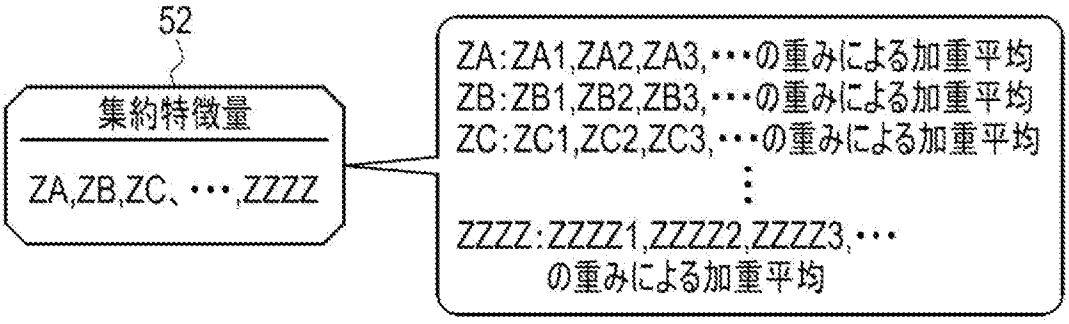
[図8]


設定補助情報
37
区分領域
3次元形状
重み
CDR
直線形状
2
ループ形状
2.5
CDR以外
直線形状
1
ループ形状
1.5
50
構造情報
No.
アミノ酸残基
区分領域
3次元形状
1
E
CH1
直線形状
2
V
CH1
直線形状
3
Q
CH1
直線形状
100
S
CDR-L1
ループ形状
101
G
CDR-L1
ループ形状
102
P
CDR-L1
ループ形状
重み 1
重み 1
重み 1
重み 2.5
重み 2.5
重み 2.5

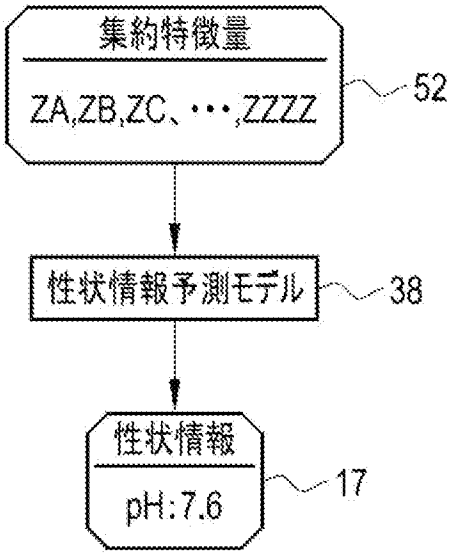
[図9]


51W
重み付き特徴量群
No.
アミノ酸残基
重み
特徴量A
特徴量B
特徴量C
・・・
特徴量ZZZ
1
E
1
ZA1
ZB1
ZC1
ZZZZ1
2
V
1
ZA2
ZB2
ZC2
ZZZZ2
3
Q
1
ZA3
ZB3
ZC3
ZZZZ3

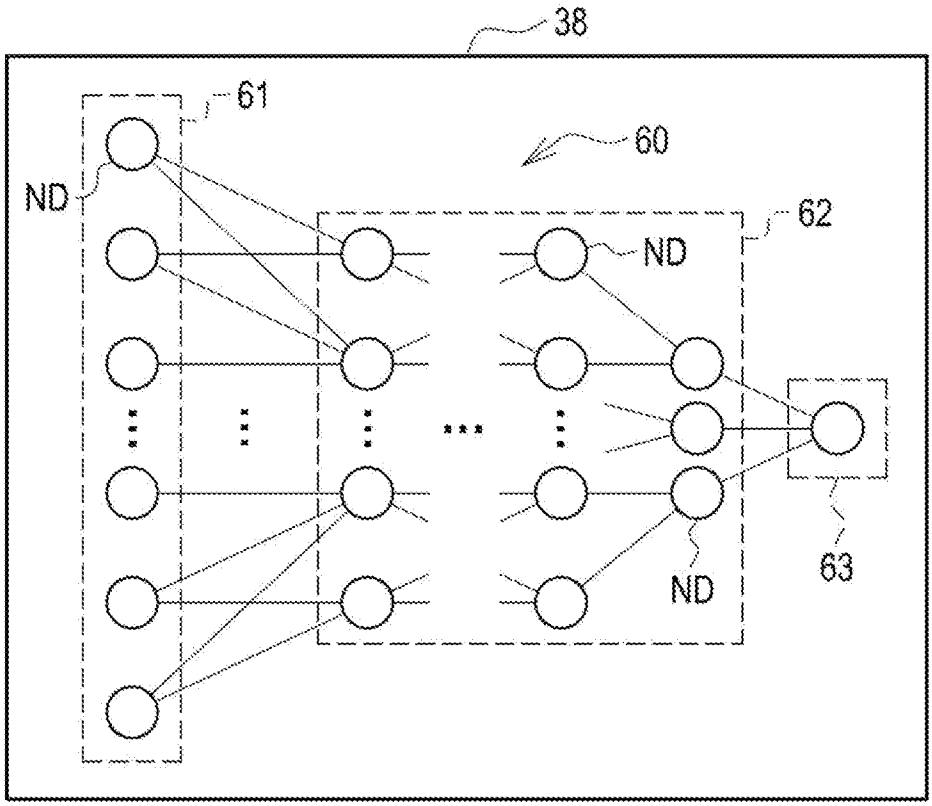
[図10]


52
集約特徴量
ZA,ZB,ZC、・・・,ZZZZ
ZA:ZA1,ZA2,ZA3,・・・の重みによる加重平均
ZB:ZB1,ZB2,ZB3,・・・の重みによる加重平均
ZC:ZC1,ZC2,ZC3,・・・の重みによる加重平均
ZZZZ:ZZZZ1,ZZZZ2,ZZZZ3,・・・
の重みによる加重平均

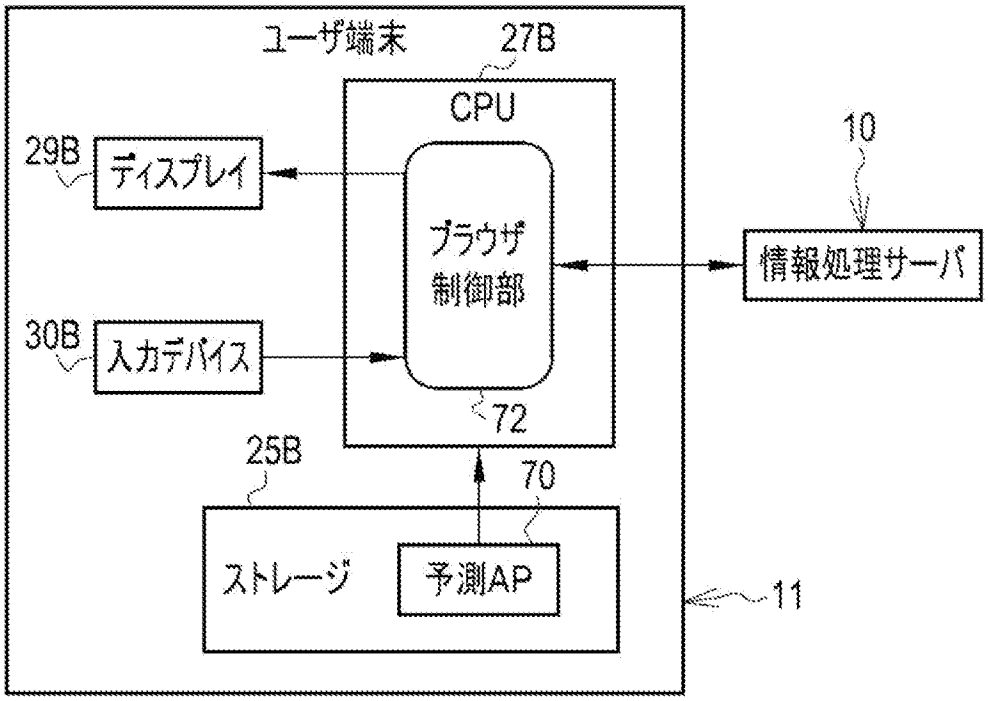
[図11]


集約特徴量
ZA,ZB,ZC、・・・,ZZZZ
52
性状情報予測モデル
38
性状情報
pH:7.6
17

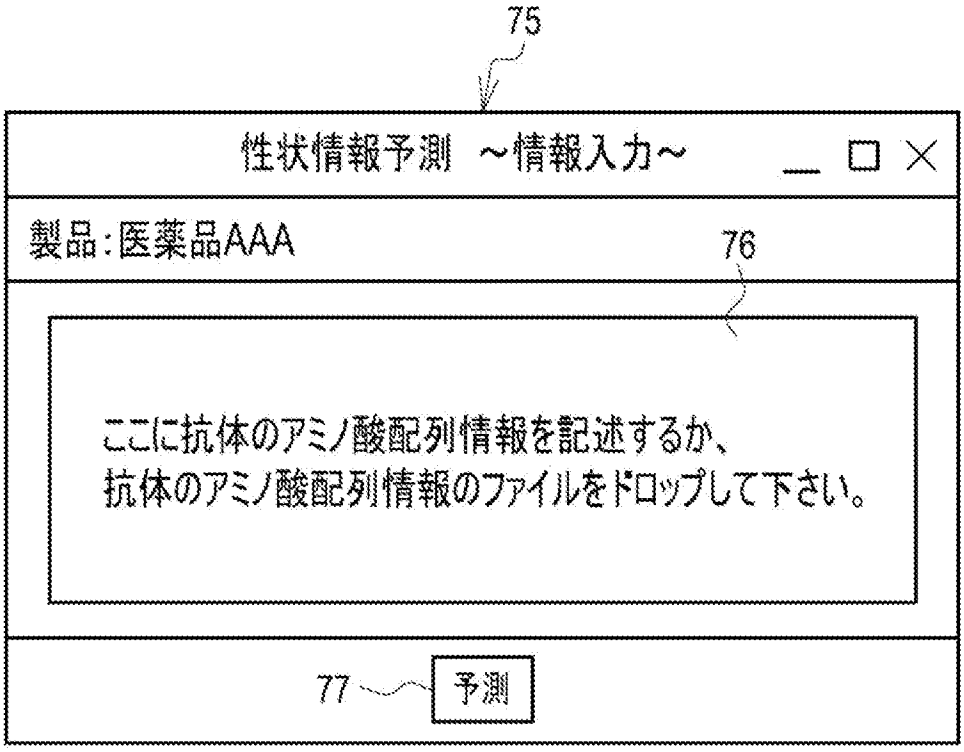
[図12]


38
61
60
ND
62
ND
ND
63

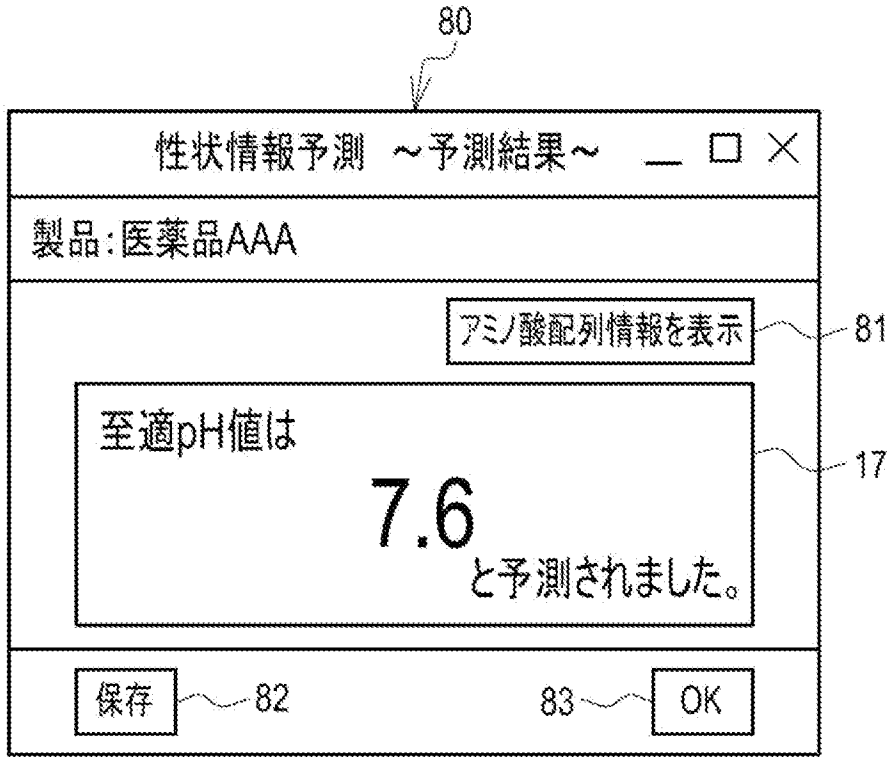
[図13]


ユーザ端末
27B
CPU
29B
ディスプレイ
10
ブラウザ
制御部
情報処理サーバ
30B
入力デバイス
72
25B
70
ストレージ
予測AP
11

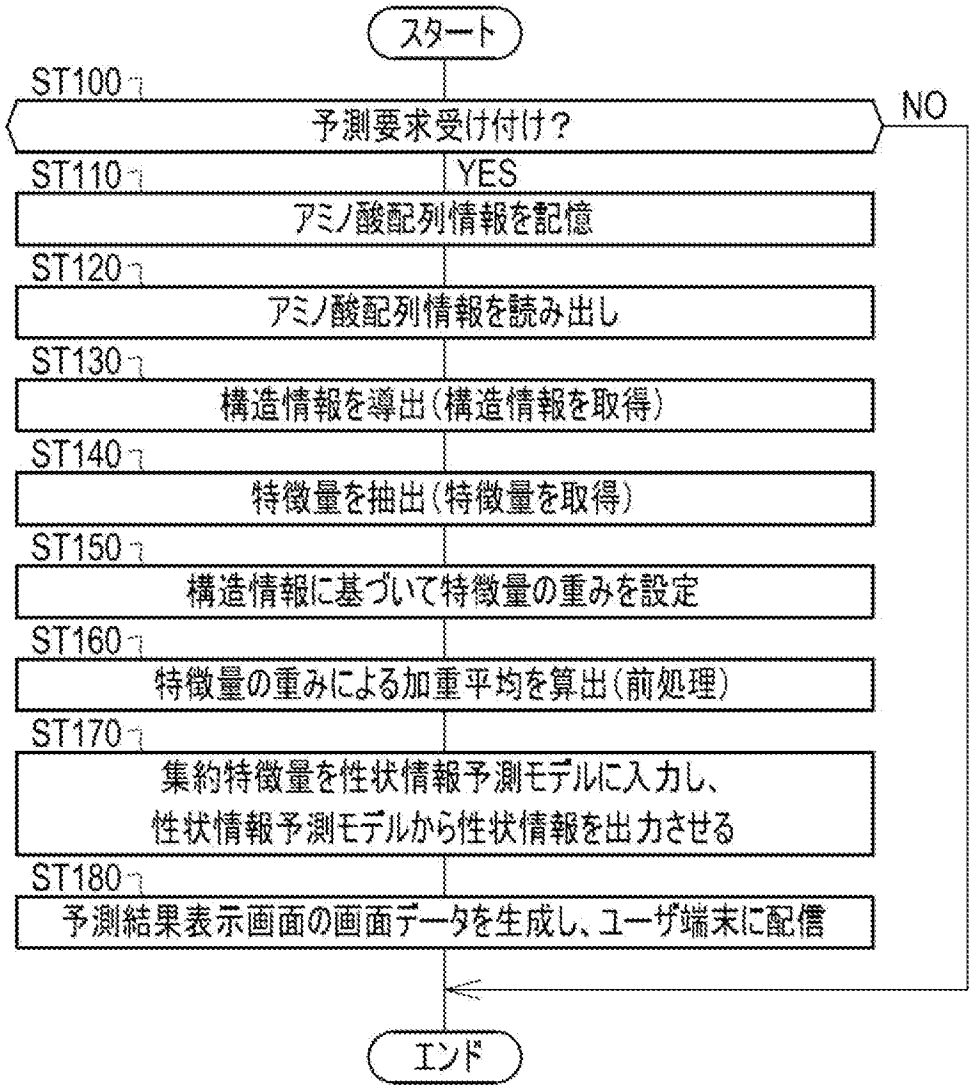
[図14]


75
性状情報予測 ～情報入力～
製品:医薬品AAA
76
ここに抗体のアミノ酸配列情報を記述するか、
抗体のアミノ酸配列情報のファイルをドロップして下さい。
77
予測

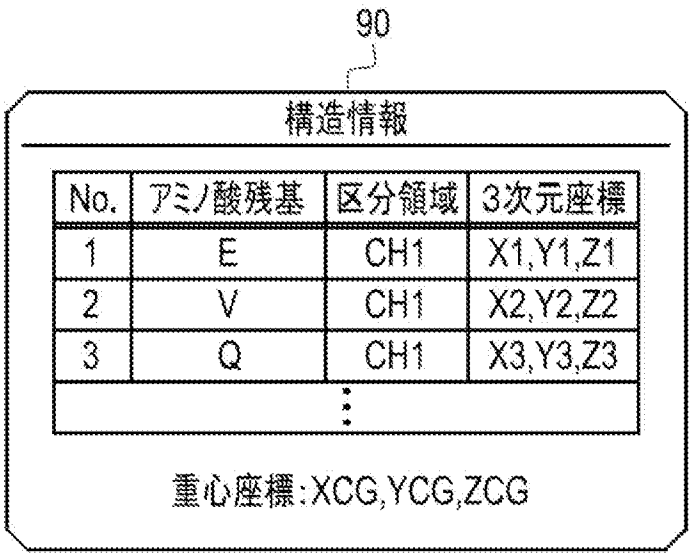
[図15]


80
性状情報予測 ～予測結果～
製品:医薬品AAA
アミノ酸配列情報を表示
81
至適pH値は
7.6
と予測されました。
17
保存
82
83
OK

[図16]


スタート
ST100
予測要求受け付け？
NO
YES
ST110
アミノ酸配列情報を記憶
ST120
アミノ酸配列情報を読み出し
ST130
構造情報を導出（構造情報を取得）
ST140
特徴量を抽出（特徴量を取得）
ST150
構造情報に基づいて特徴量の重みを設定
ST160
特徴量の重みによる加重平均を算出（前処理）
ST170
集約特徴量を性状情報予測モデルに入力し、
性状情報予測モデルから性状情報を出力させる
ST180
予測結果表示画面の画面データを生成し、ユーザ端末に配信
エンド

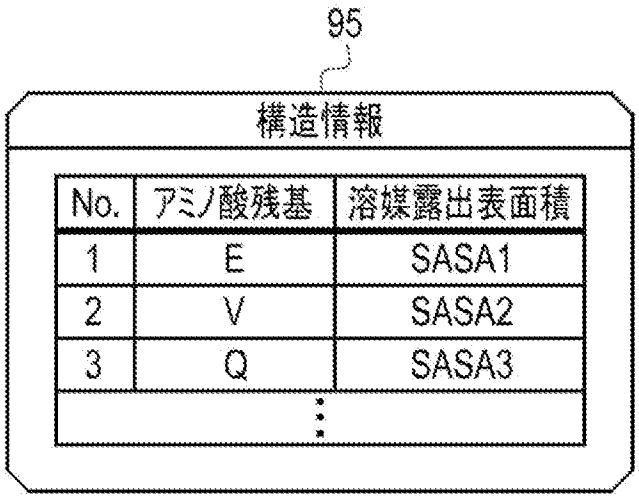
[図17]


90
構造情報
No.
アミノ酸残基
区分領域
3次元座標
1
E
CH1
X1,Y1,Z1
2
V
CH1
X2,Y2,Z2
3
Q
CH1
X3,Y3,Z3
重心座標：XCG,YCG,ZCG

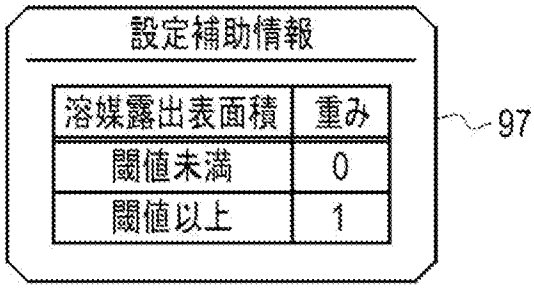
[図18]


設定補助情報
区分領域
重心からの距離
重み
CDR
閾値未満
2
閾値以上
2.5
CDR以外
閾値未満
1
閾値以上
1.5
92

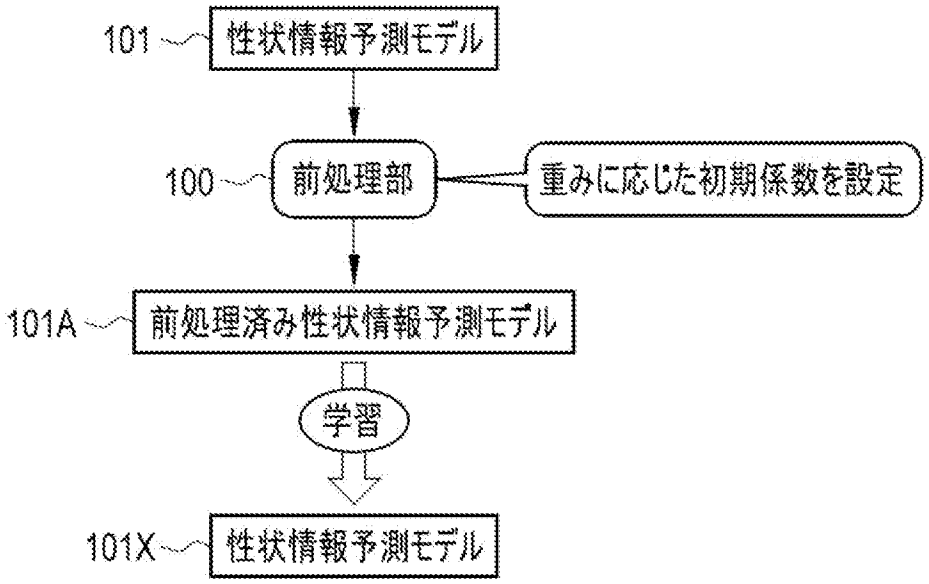
[図19]


95
構造情報
No.
アミノ酸残基
溶媒露出表面積
1
E
SASA1
2
V
SASA2
3
Q
SASA3

[図20]


設定補助情報
溶媒露出表面積
重み
閾値未満
0
閾値以上
1
97

[図21]


101
性状情報予測モデル
100
前処理部
重みに応じた初期係数を設定
101A
前処理済み性状情報予測モデル
学習
101X
性状情報予測モデル

INTERNATIONAL SEARCH REPORT

International application No.
PCT/JP2024/032771

A. CLASSIFICATION OF SUBJECT MATTER

G16B 30/00(2019.01)i; ***G16B 40/00***(2019.01)i
FI: G16B30/00; G16B40/00

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G06Q10/00-99/00; G16B5/00-99/00; G16C10/00-99/00

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Published examined utility model applications of Japan 1922-1996
Published unexamined utility model applications of Japan 1971-2024
Registered utility model specifications of Japan 1996-2024
Published registered utility model applications of Japan 1994-2024

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	齋藤 裕, 深層学習によるタンパク質の機能予測と設計, 生物工学会誌 [online], vol. 100, no. 11, 公益社団法人 日本生物工学会, 2022, pp. 589-592, [retrieved on 30 September 2024], Internet: <URL: https://www.jstage.jst.go.jp/article/seibutsukogaku/100/11/100_100.11_589/_pdf>, Online ISSN: 2435-8630, Print ISSN: 0919-3758, <DOI: 10.34565/seibutsukogaku.100.11_589>, (SAITO, Yutaka, Biotechnology, The Society for Biotechnology, Japan), non-official translation (Protein function prediction and design by deep learning) p. 589 "Basic structure of a deep learning model", right column, lines 14-22, p. 590, left column, lines 25-31	1-2, 4-5, 7, 9-13
A	entire text, all drawings	3, 6, 8
A	JP 2022-527381 A (ETH ZUERICH) 01 June 2022 (2022-06-01) entire text, all drawings	1-13
A	JP 2020-77206 A (MIRACA RES INSTITUTE GK) 21 May 2020 (2020-05-21) entire text, all drawings	1-13

☐ Further documents are listed in the continuation of Box C. ☑ See patent family annex.

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"D" document cited by the applicant in the international application
"E" earlier application or patent but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
30 September 2024	**15 October 2024**
Name and mailing address of the ISA/JP **Japan Patent Office (ISA/JP)** **3-4-3 Kasumigaseki, Chiyoda-ku, Tokyo 100-8915** **Japan**	Authorized officer Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/JP2024/032771

Patent document cited in search report	Publication date (day/month/year)	Patent family member(s)	Publication date (day/month/year)
JP 2022-527381 A	01 June 2022	US 2022/0157403 A1 entire text, all drawings WO 2020/208555 A1 EP 3953943 A1 CN 113853656 A	
JP 2020-77206 A	21 May 2020	(Family: none)	

国際調査報告

国際出願番号
PCT/JP2024/032771

A.　発明の属する分野の分類（国際特許分類（ＩＰＣ））

G16B 30/00(2019.01)i; G16B 40/00(2019.01)i
FI: G16B30/00; G16B40/00

B.　調査を行った分野

調査を行った最小限資料（国際特許分類（ＩＰＣ））

G06Q10/00-99/00; G16B5/00-99/00; G16C10/00-99/00

最小限資料以外の資料で調査を行った分野に含まれるもの

日本国実用新案公報	１９２２－１９９６年
日本国公開実用新案公報	１９７１－２０２４年
日本国実用新案登録公報	１９９６－２０２４年
日本国登録実用新案公報	１９９４－２０２４年

国際調査で使用した電子データベース（データベースの名称、調査に使用した用語）

C.　関連すると認められる文献

引用文献のカテゴリー*	引用文献名　及び一部の箇所が関連するときは、その関連する箇所の表示	関連する請求項の番号
X	齋藤 裕, 深層学習によるタンパク質の機能予測と設計, 生物工学会誌 [online], 第100巻, 第11号, 公益社団法人 日本生物工学会, 2022, p. 589-592, [検索日 2024.09.30], インターネット: <URL: https://www.jstage.jst.go.jp/article/seibutsukogaku/100/11/100_100.11_589/_pdf>, Online ISSN: 2435-8630, Print ISSN: 0919-3758, <DOI: 10.34565/seibutsukogaku.100.11_589> p. 589“深層学習モデルの基本構造”, 右欄14-22行目及びp. 590の左欄25-31行目	1-2,4-5,7,9-13
A	全文, 全図	3,6,8
A	JP 2022-527381 A (エーテーハー チューリッヒ) 01.06.2022 (2022-06-01) 全文, 全図	1-13
A	JP 2020-77206 A (合同会社みらか中央研究所) 21.05.2020 (2020-05-21) 全文, 全図	1-13

☐ Ｃ欄の続きにも文献が列挙されている。

☑ パテントファミリーに関する別紙を参照。

*　引用文献のカテゴリー
“A” 特に関連のある文献ではなく、一般的技術水準を示すもの
“D” 国際出願で出願人が先行技術文献として記載した文献
“E” 国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの
“L” 優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献（理由を付す）
“O” 口頭による開示、使用、展示等に言及する文献
“P” 国際出願日前で、かつ優先権の主張の基礎となる出願の日の後に公表された文献
“T” 国際出願日又は優先日後に公表された文献であって出願と抵触するものではなく、発明の原理又は理論の理解のために引用するもの
“X” 特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの
“Y” 特に関連のある文献であって、当該文献と他の１以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの
“&” 同一パテントファミリー文献

国際調査を完了した日	国際調査報告の発送日
30.09.2024	15.10.2024

名称及びあて先

日本国特許庁(ISA/JP)
〒100-8915
日本国
東京都千代田区霞が関三丁目４番３号

権限のある職員（特許庁審査官）

前田　侑香 5R 2569

電話番号 03-3581-1101 内線 3560

国際調査報告
パテントファミリーに関する情報

国際出願番号
PCT/JP2024/032771

引用文献	公表日	パテントファミリー文献	公表日
JP 2022-527381 A	01.06.2022	US 2022/0157403 A1 全文，全図 WO 2020/208555 A1 EP 3953943 A1 CN 113853656 A	
JP 2020-77206 A	21.05.2020	（ファミリーなし）	