

(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
2022年6月2日(02.06.2022)



(10) 国際公開番号

WO 2022/113338 A1

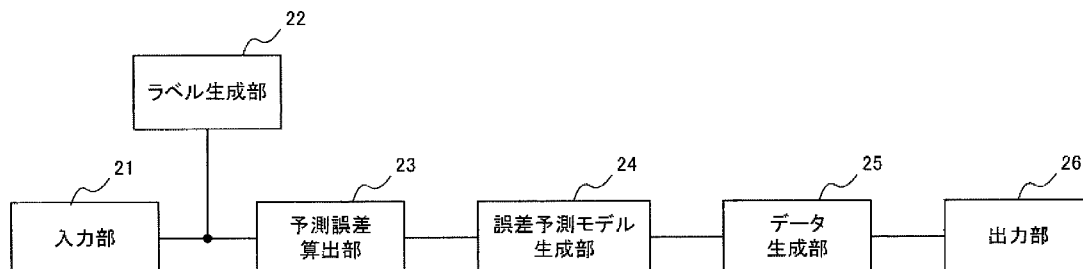
- (51) 国際特許分類:
G06N 20/00 (2019.01)
- (21) 国際出願番号: PCT/JP2020/044483
- (22) 国際出願日: 2020年11月30日(30.11.2020)
- (25) 国際出願の言語: 日本語
- (26) 国際公開の言語: 日本語
- (71) 出願人: 日本電気株式会社 (NEC CORPORATION) [JP/JP]; 〒1088001 東京都港区芝五丁目7番1号 Tokyo (JP).
- (72) 発明者: 畠山 優太 (HATAKEYAMA, Yuta); 〒1088001 東京都港区芝五丁目7番1号 日本電気株式会社内 Tokyo (JP). 岡嶋 穰 (OKAJIMA, Yuzuru); 〒1088001 東京都港区芝五丁目7番1号 日本電気株式会社内 Tokyo (JP).
- (74) 代理人: 中村 聡延, 外 (NAKAMURA, Toshinobu et al.); 〒1040031 東京都中央区京橋一丁目16番10号 オークビル京橋3階 東京セントラル特許事務所内 Tokyo (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, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, IT, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ,

(54) Title: INFORMATION PROCESSING DEVICE, INFORMATION PROCESSING METHOD, AND RECORDING MEDIUM

(54) 発明の名称: 情報処理装置、情報処理方法、及び、記録媒体

[図4]

100



- 21 Input unit
22 Label generation unit
23 Prediction error calculation unit
24 Error prediction model generation unit
25 Data generation unit
26 Output unit

(57) Abstract: An information processing device, wherein an input means accepts a training example composed of feature quantities. A label generation means adds a label to the training example. An error calculation means generates one or more student models using the training example to which the label has been added, and calculates errors between the label and predictions by the student models. An error prediction model generation means generates an error prediction model that predicts errors. An output means outputs an example for which it is predicted, on the basis of the error prediction model, that errors will increase.

(57) 要約: 情報処理装置において、入力手段は、特徴量からなる訓練用例を受け付ける。ラベル生成手段は、訓練用例にラベルを付与する。誤差算出手段は、ラベルが付与された訓練用例を用いて1つ以上の生徒モデルを生成し、当該生徒モデルによる予測とラベルとの誤差を算出する。誤差予測モデル生成手段は、誤差を予測するモデルである誤差予測モデルを生成する。出力手段は、誤差予測モデルに基づき、誤差が大きくなると予測される用例を出力する。

[続葉有]

WO 2022/113338 A1

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, GH, GM, KE, LR, LS,
MW, MZ, NA, RW, 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, 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))

明 細 書

発明の名称：情報処理装置、情報処理方法、及び、記録媒体

技術分野

[0001] 本発明は、機械学習モデルの精度を向上させるための技術に関する。

背景技術

[0002] 教師あり学習により機械学習モデルの精度を向上させる技術として、能動学習が知られている。能動学習とは、現在の機械学習モデルではうまく予測できない用例に対して教師（オラクル）がラベルを付与して用例を生成し、その用例を用いて機械学習モデルを再訓練することによりモデルの精度を向上させる手法である。

[0003] 能動学習の手法は、基本的に「生徒モデルが曖昧な予測や矛盾する予測を出力する用例」をうまく予測できない用例とみなし、その用例に対してラベルを付与して再訓練を行う。能動学習の例としては、Uncertainty sampling やQuery-by-committee (QBC) などが知られている。Uncertainty sampling は生徒モデルが作る決定境界に近い用例に対してラベルを付与する手法であり、Query-by-committeeは複数の生徒モデルが矛盾する答えを出す用例に対してラベルを付与する手法である。

[0004] また、非特許文献1は、GAN (Generative Adversarial Network) と能動学習を組み合わせた手法を提案している。この手法では、GANを使って、対象となる分類器が曖昧な予測を出力する人工用例を作成している。

先行技術文献

特許文献

[0005] 非特許文献1：Kong, Q., Tong, B., Klinkigt, M., Watanabe, Y., Akira, N., and Murakami, T. Active generative adversarial network for image classification. In Association for the Advancement of Artificial Intelligence, 2019. Airxiv: <https://arxiv.org/abs/1906.07133>

発明の概要

発明が解決しようとする課題

[0006] しかし、生徒モデルが曖昧な予測を出力することと、生徒モデルが予測を間違ふこととはイコールではない。例えば、決定境界から遠くても生徒モデルの予測が間違いであることがある。また、生徒モデルがほぼ信頼度「1」で予測したとしても、実際には予測が間違っていることもある。これは特に生徒モデルの予測が信頼できない場合に顕著である。よって、上記の能動学習の手法では、予測精度を大きく向上させる用例を効率的に発見することが難しい。

[0007] 本発明の1つの目的は、予測精度を大きく向上させる用例を効率的に発見することにある。

課題を解決するための手段

[0008] 本発明の一つの観点は、情報処理装置であって、
特徴量からなる訓練用例を受け付ける入力手段と、
前記訓練用例にラベルを付与するラベル生成手段と、
ラベルが付与された前記訓練用例を用いて1つ以上の生徒モデルを生成し、
当該生徒モデルによる予測と前記ラベルとの誤差を算出する誤差算出手段と、
前記誤差を予測するモデルである誤差予測モデルを生成する誤差予測モデル生成手段と、
前記誤差予測モデルに基づき、誤差が大きくなると予測される用例を出力する出力手段と、
を備える。

[0009] 本発明の他の観点は、情報処理方法であって、
特徴量からなる訓練用例を受け付け、
前記訓練用例にラベルを付与し、
ラベルが付与された前記訓練用例を用いて1つ以上の生徒モデルを生成し、
当該生徒モデルによる予測と前記ラベルとの誤差を算出し、
前記誤差を予測するモデルである誤差予測モデルを生成し、

前記誤差予測モデルに基づき、誤差が大きくなると予測される用例を出力する。

- [0010] 本発明の他の観点は、記録媒体であって、
特徴量からなる訓練用例を受け付け、
前記訓練用例にラベルを付与し、
ラベルが付与された前記訓練用例を用いて1つ以上の生徒モデルを生成し、
当該生徒モデルによる予測と前記ラベルとの誤差を算出し、
前記誤差を予測するモデルである誤差予測モデルを生成し、
前記誤差予測モデルに基づき、誤差が大きくなると予測される用例を出力する処理をコンピュータに実行させるプログラムを記録する。

図面の簡単な説明

- [0011] [図1]実施形態の手法を概念的に示す図である。
[図2]実施形態の情報処理装置を概念的に示す図である。
[図3]第1実施形態の情報処理装置のハードウェア構成を示す図である。
[図4]第1実施形態の情報処理装置の機能構成を示す図である。
[図5]誤差予測モデルを生成する方法を模式的に示す。
[図6]予測誤差の複数の観測点に基づきガウス過程回帰を行った結果の例を示す。
[図7]予測誤差の過学習を説明する図である。
[図8]訓練用例から誤差算出用例を生成する具体例を示す。
[図9]第1実施形態の情報処理装置による処理のフローチャートである。
[図10]第2実施形態に係る情報処理装置の機能構成を示す図である。
[図11]第2実施形態の情報処理装置による処理のフローチャートである。

発明を実施するための形態

- [0012] 以下、図面を参照して、本発明の好適な実施形態について説明する。

<第1実施形態>

[基本原理]

既知の能動学習の手法では、生徒モデルが曖昧な予測を出力する用例に対

してラベルを付与して再訓練を行っている。しかし、前述のように、生徒モデルが曖昧な予測を出力することと、生徒モデルが予測を間違えることとはイコールではなく、生徒モデルが信頼度「1」で出力した予測が間違っていることがある。これは、生徒モデルのみに基づいて、再訓練に用いる用例を選択している点に原因がある。即ち、生徒モデルの予測を、その生徒モデル自身が出力する信頼度や確率に基づいて評価しているため、再訓練用に選択される用例の是非は、その生徒モデルの実際の精度に左右されてしまう。

[0013] そこで、本実施形態では、絶対的に正しい予測を出力するとみなすことができる教師モデルを用意し、生徒モデルの予測を、その教師モデルの予測と比較することにより評価する。これにより、生徒モデルの予測が教師モデルの予測に近ければ、生徒モデルの予測は信頼できると考えられる。一方、生徒モデルの予測が教師モデルの予測と遠ければ、生徒モデルの予測は疑わしいと考えられる。よって、生徒モデルの予測と教師モデルの予測との誤差が大きくなる用例を再訓練用の用例として選択すれば、精度向上に大きく寄与する用例を得ることが可能となる。

[0014] 図1は、本実施形態の手法を概念的に示す図である。前述のように、訓練の対象となる生徒モデルに加えて、上述の絶対的に正しい予測を出力するとみなすことができる教師モデルを用意する。まず、複数の訓練用例について、生徒モデルと教師モデルのそれぞれで予測を行い、予測の誤差を計算する。次に、計算された誤差に基づいて、生徒モデルと教師モデルの予測の誤差を推定するモデル（以下、「誤差予測モデル」と呼ぶ。）を生成する。そして、この誤差予測モデルを用いて再訓練用の用例を選択する。これにより、生徒モデルの精度向上に寄与する用例を選択することが可能となる。

[0015] [情報処理装置の全体構成]

図2は、本実施形態の情報処理装置を概念的に示す図である。情報処理装置100には、複数のラベルなし訓練用例が入力される。情報処理装置100は、まず、入力されたラベルなし訓練用例に対して、上述の教師モデルでラベル付けを行う。このラベルは、教師モデルによる予測に相当する。次に

、情報処理装置100は、ラベル付けした訓練用例を用いて生徒モデルを生成する。次に、情報処理装置100は、生成した生徒モデルにより予測を行い、教師モデルが付与したラベルと、生徒モデルの予測との誤差を示す誤差予測モデルを生成する。こうして生成した誤差予測モデルは、用例ごとに、その用例に対する生徒モデルと教師モデルの予測の誤差を示すモデルとなる。そして、情報処理装置100は、誤差予測モデルを用いて、誤差が大きくなると予測される用例を出力する。この用例に対してラベルを付与して生徒モデルの再訓練を行うことにより、生徒モデルの精度向上が期待できる。

[0016] [ハードウェア構成]

図3は、第1実施形態の情報処理装置100のハードウェア構成を示すブロック図である。図示のように、情報処理装置100は、入力IF (Interface) 11と、プロセッサ12と、メモリ13と、記録媒体14と、データベース(DB) 15と、を備える。

[0017] 入力IF 11は、データの入出力を行う。具体的に、入力IF 11は、特徴量からなる訓練用例を取得するとともに、誤差予測モデルに基づき、誤差が大きくなると予測される用例を出力する。

[0018] プロセッサ12は、CPU (Central Processing Unit)、GPU (Graphics Processing Unit) などのコンピュータであり、予め用意されたプログラムを実行することにより、情報処理装置100の全体を制御する。特に、プロセッサ12は、誤差予測モデルに基づき、誤差が大きくなると予測される用例を出力する用例出力処理を行う。

[0019] メモリ13は、ROM (Read Only Memory)、RAM (Random Access Memory) などにより構成される。メモリ13は、プロセッサ12により実行される各種のプログラムを記憶する。また、メモリ13は、プロセッサ12による各種の処理の実行中に作業メモリとしても使用される。

[0020] 記録媒体14は、ディスク状記録媒体、半導体メモリなどの不揮発性で非

一時的な記録媒体であり、情報処理装置 100 に対して着脱可能に構成される。記録媒体 14 は、プロセッサ 12 が実行する各種のプログラムを記録している。

[0021] DB 15 は、入力 IF 11 から入力される訓練用例を記憶する。また、DB 15 には、訓練用例を用いて生成された誤差予測モデルの情報などが記憶される。

[0022] [機能構成]

図 4 は、情報処理装置 100 の機能構成を示すブロック図である。情報処理装置 100 は、入力部 21 と、ラベル生成部 22 と、予測誤差算出部 23 と、誤差予測モデル生成部 24 と、データ生成部 25 と、出力部 26 とを備える。

[0023] 入力部 21 には、生徒モデルの訓練に使用するラベルなし訓練用例と、訓練済み教師モデルとが入力される。訓練用例は、複数次元の特徴量により構成される。入力部 21 は、ラベルなし訓練用例と訓練済み教師モデルとをラベル生成部 22 へ出力する。また、入力部 21 は、ラベルなし訓練用例を予測誤差算出部 23 へ出力する。

[0024] ラベル生成部 22 は、訓練済み教師モデルを用いて、入力されたラベルなし訓練用例に対するラベルを生成し、予測誤差算出部 23 へ出力する。なお、このラベルは、入力されたラベルなし訓練用例に対する教師モデルの予測に相当する。

[0025] 予測誤差算出部 23 は、入力部 21 からラベルなし訓練用例を取得するとともに、ラベル生成部 22 から、各訓練用例に対して付与されたラベルを取得する。これにより、予測誤差算出部 23 には、ラベル付き訓練用例が用意される。予測誤差算出部 23 は、このラベル付き訓練用例を用いて生徒モデルを訓練し、訓練済みの生徒モデルを生成する。

[0026] 次に、予測誤差算出部 23 は、生成した生徒モデルを用いて予測を行う。そして、予測誤差算出部 23 は、生徒モデルによる予測と、ラベル生成部 22 から入力されたラベルとの誤差、即ち、生徒モデルによる予測と教師モデ

ルによる予測との誤差（以下、「予測誤差」と呼ぶ。）を算出し、誤差予測モデル生成部24へ出力する。なお、後に詳しく説明するが、この予測誤差の算出においては、生徒モデルの訓練に使用した訓練用例とは異なる用例が用いられる。

[0027] 誤差予測モデル生成部24は、予測誤差算出部23から、複数の用例に対する予測誤差を取得し、誤差予測モデルを生成する。誤差予測モデルは、前述のように、個々の用例について、教師モデルと生徒モデルの予測誤差を推定するモデルである。ここで、誤差予測モデル生成部24は、誤差予測モデルとして微分可能なモデルを生成する。これは、誤差予測モデルを微分可能なモデルとすることにより、生徒モデルが微分不可なモデルであっても、誤差が大きくなる用例を効率的に探索できるからである。好適な例では、誤差予測モデルとして回帰モデルが使用される。例えば、生徒モデルが決定木のような微分不可なモデルの場合でも、誤差予測モデルとして微分可能な回帰モデルを使用すれば、回帰モデルの勾配を計算することにより、予測誤差が大きくなる用例を効率的に発見することが可能となる。

[0028] 図5は、誤差予測モデルを生成する方法を模式的に示す。複数の用例について、教師モデルと生徒モデルの予測誤差 $E_1 \sim E_3$ が得られたとする。誤差予測モデル生成部24は、各予測誤差 $E_1 \sim E_3$ を、図5に示すように横軸が用例を示し、縦軸が予測誤差を示す領域にプロットする。なお、横軸における各用例の位置は、各用例の特徴量に基づいて決定される。また、縦軸の上方ほど予測誤差の値が大きいものとする。図示のように各用例の予測誤差は離散的であり、誤差予測モデル生成部24は、予測誤差 $E_1 \sim E_3$ を通る微分可能な関数を誤差予測モデル F として生成する。

[0029] 以下、好適な例として、ガウス過程回帰を用いて誤差予測モデルを生成する例を示す。ガウス過程回帰とは、観測点 $f(x_1), \dots, f(x_n)$ から連続関数 f を予測する手法であり、より正確には、離散観測点から関数 f の確率分布をガウス分布の形で求める手法である。図6は、予測誤差算出部23が出力した予測誤差の複数の観測点 $f(x_n)$ に基づいてガウス過程回帰

を行った結果の例を示す。グラフ f は真の関数を示すが、実際には不明である。観測点 $f(x_n)$ に基づくガウス過程回帰の結果として、グラフ M で示す誤差の平均と、グレーの領域 D で示す誤差の分散とが得られる。即ち、誤差予測モデル生成部 24 は、誤差予測モデルとして、誤差の平均を示すグラフ M と、誤差の分散を示す領域 D とを生成する。このように、ガウス過程回帰は、微分可能（即ち、連続）な平均や分散を出力するので、この平均や分散を組み込んだ関数も微分可能となり、最適化が容易となる。好適な例では、誤差予測モデル生成部 24 は、生徒モデルと教師モデルの予測誤差を観測点とし、ガウス過程回帰を用いて誤差予測モデルを生成する。誤差予測モデル生成部 24 は、生成した誤差予測モデルをデータ生成部 25 へ出力する。

[0030] データ生成部 25 は、誤差予測モデルに基づいて、誤差が大きい場所を補完する用例を生成し、出力部 26 へ出力する。例えば、誤差予測モデルとして図 5 に示す関数 F が生成された場合、データ生成部 25 は、関数 F 上の誤差が大きい点 p に対応する用例 P を生成する。具体的に、入力部 21 に入力された複数の訓練用例中に用例 P に相当する訓練用例が存在する場合、データ生成部 25 はその訓練用例を出力する。また、用例 P に相当する訓練用例が存在しない場合、データ生成部 25 は、それに近い訓練用例の特徴量に基づいて用例 P に相当する特徴量を人工的に生成し、人工用例 P として出力する。なお、データ生成部 25 は、誤差予測モデルに基づいて、誤差が最大となる点に対応する用例を出力してもよく、誤差が大きい方から所定数の点に対応する用例を出力してもよく、誤差が所定の閾値以上である点に対応する用例を出力してもよい。

[0031] また、図 6 に例示するように、誤差予測モデル生成部 24 がガウス過程回帰を用い、誤差予測モデルとして誤差の平均 M と分散 D を出力した場合、データ生成部 25 は、誤差の平均 M と分散 D の少なくとも一方が大きい点に対応する用例を、誤差が大きくなる予測される用例として出力する。具体的には、データ生成部 25 は、誤差の平均 M や分散 D が最大値に近い所定数の点に対応する用例、又は、誤差の平均 M や分散 D が所定の閾値より大きい複数

の点に対応する用例などを出力してもよい。また、データ生成部 25 は、ガウス過程回帰を利用したベイズ最適化などにより、真の関数 f の最大点を予測し、最大点に対応する用例を、誤差が大きくなると予測される用例として出力してもよい。

[0032] 出力部 26 は、データ生成部 25 から入力された用例を、誤差が大きくなると予測される用例として出力する。こうして出力された用例は、生徒モデルの再訓練に使用される。具体的には、出力された用例に対して、ラベル生成部 22 で使用した教師モデルを用いてラベルを付与して訓練用例とし、生徒モデルの訓練に使用すればよい。もしくは、出力された用例に対して、ラベル生成部 22 で使用した教師モデルとは別の教師モデル又は人手によりラベルを付与してもよい。

[0033] [予測誤差の過学習]

次に、予測誤差の過学習について説明する。上記の情報処理装置 100 では、入力部 21 から入力されるラベルなし訓練用例に対して教師モデルを用いてラベルを付与してラベル付き訓練用例を生成し、そのラベル付き訓練用例を用いて生徒モデルを訓練している。よって、予測誤差算出部 23 が、生徒モデルの訓練に使用した訓練用例と同じ用例を用いて教師モデルと生徒モデルの予測誤差を算出すると、各用例における教師モデルと生徒モデルの予測は一致するため、予測誤差がゼロになってしまう、よって、この予測誤差を用いて誤差予測モデル生成部 24 が誤差予測モデルを生成すると、誤差予測モデルは訓練用例に対応する全ての点で誤差がゼロになってしまう。つまり、誤差予測モデルを用いた予測誤差の予測自体に過学習が発生してしまい、本来の誤差よりも小さい誤差が予測されてしまう。これを「予測誤差の過学習」と呼ぶ。

[0034] 図 7 は、予測誤差の過学習を説明する図である。図 7 において、複数の点 71 が訓練用例を示し、実線のグラフ 72 が教師モデルを示すとする。生徒モデルは、訓練用例に対して教師モデルが付与したラベルを教師データとして訓練されるため、破線のグラフ 73 で示すように、訓練用例 71 の位置で

は教師モデルとの予測誤差がゼロとなるように訓練される。よって、生徒モデルの訓練に使用した訓練用例と同じ用例を用いて教師モデルと生徒モデルの予測誤差を算出すると、誤差予測モデルの生成に使用される予測誤差は全てゼロとなり、汎化性を有する誤差予測モデルを生成することができなくなってしまう。

[0035] そこで、本実施形態では、生徒モデルの訓練に用いた訓練用例とは異なる用例を用いて、誤差予測モデルの生成に用いる予測誤差を算出する。即ち、生徒モデルの訓練に用いた訓練用例（以下、「生徒モデル訓練用例」と呼ぶ。）と、誤差予測モデルの生成に使用する予測誤差を算出するための用例（以下、「誤差算出用例」と呼ぶ。）を異なる用例とする。以下、生徒モデル訓練用例と異なる誤差算出用例を生成する方法を説明する。

[0036] （方法1）オーバーサンプリングにより誤差算出用例を生成する。

オーバーサンプリングとは、人工的に用例を生成する手法であり、例えば、SMOTE、MUNGEなどが挙げられる。具体的には、予め用意された訓練用例の全てを生徒モデル訓練用例として用い、生徒モデルを訓練する。また、オーバーサンプリングにより訓練用例からラベルなし用例 x' を新たに作成し、誤差算出用例とする。そして、新たなラベルなし用例 x' を用いて、例えば下記の式により、教師モデルと生徒モデルの予測誤差を算出する。

$$\text{予測誤差} = |\text{教師.predict}(x') - \text{生徒.predict}(x')|$$

[0037] （方法2）訓練用例を分割して誤差算出用例を生成する。

方法2では、教師モデルによってラベルを付与された訓練用例を分割し、分割した一部の訓練用例を生徒モデル訓練用例として用いて生徒モデルを訓練する。また、分割した残りの訓練用例を誤差算出用例として用い、教師モデルと生徒モデルの予測誤差を算出する。

[0038] 図8は、方法2により訓練用例を分割して誤差算出用例を生成する具体例を示す。図示のように、入力部21からN個（図8の例では $N=5$ ）の訓練用例からなるデータセットが入力されるとする。まず、ラベル生成部22は

、訓練用例の全データに教師モデルを用いてラベルを付与する（工程 P 1）。

[0039] 次に、予測誤差算出部 2 3 は、訓練用例に対して重複ありランダムサンプリング（ブートストラップサンプリング）を行い、M 個（図 8 の例では M = 3）のブートストラップサンプル群を生成する（工程 P 2）。各ブートストラップサンプル群のデータ数は N 個である。そして、予測誤差算出部 2 3 は、各ブートストラップサンプル群を用いて生徒モデルを作成する（工程 P 3）。即ち、各ブートストラップサンプル群は、生徒モデル訓練用例として使用される。これにより、M 個の生徒モデルが生成される。

[0040] 各ブートストラップサンプル群は、訓練用例を重複ありランダムサンプリングして生成されているため、訓練用例に含まれるが各ブートストラップサンプル群には選ばれなかったサンプルが存在する。これを O O B（Out-of-Bag）と呼ぶ。O O B は、ブートストラップサンプル群には含まれていないため、生徒モデルの生成には使用されていない。そこで、予測誤差算出部 2 3 は、各ブートストラップサンプル群の O O B を誤差算出用例として使用し、教師モデルと生徒モデルの予測誤差を算出する。具体的には、予測誤差算出部 2 3 は、各ブートストラップサンプル群の O O B について、そのブートストラップサンプル群に対応する生徒モデルと、教師モデルとによる予測を取得する。そして、予測誤差算出部 2 3 は、下記の式により、M 個のブートストラップサンプル群の各々についての予測誤差を算出し、その平均を予測誤差として誤差予測モデル生成部 2 4 へ出力する（工程 P 4）。

$$\text{予測誤差} = | \text{教師.predict(OOB)} - \text{生徒.predict(OOB)} |$$

こうして、生徒モデルの生成に使用した訓練用例と異なる用例を用いて、教師モデルと生徒モデルの予測誤差を算出することができる。

[0041] （方法 3）別の訓練用例を取得する。

方法 3 では、入力部 2 1 に入力された全ての訓練用例を生徒モデル訓練用例として使用し、生徒モデルを生成する。一方、その訓練用例と異なるラベルなし用例を別途取得し、誤差算出用例として使用する。

[0042] [情報処理装置による処理]

次に、情報処理装置100により、誤差の大きいと予測される用例を出力する処理（以下、「用例出力処理」とも呼ぶ。）について説明する。図9は、用例出力処理のフローチャートである。この処理は、図3に示すプロセッサ12が予め用意されたプログラムを実行し、図4に示す各要素として動作することにより実現される。

[0043] まず、入力部21は、ラベルなし訓練用例と教師モデルを取得する（ステップS11）。次に、ラベル生成部22は、教師モデルを用いて、ラベルなし訓練用例にラベルを付与する（ステップS12）。次に、予測誤差算出部23は、ステップS12でラベルが付与された訓練用例を用いて生徒モデルを生成する（ステップS13）。

[0044] 次に、予測誤差算出部23は、教師モデルと生徒モデルの予測誤差を算出する（ステップS14）。この際、予測誤差算出部23は、前述のように、ステップS13において生徒モデルの生成に使用された生徒モデル訓練用例とは異なる誤差算出用例について、教師モデルと生徒モデルの予測誤差を算出する。

[0045] 次に、誤差予測モデル生成部24は、算出された予測誤差を用いて、例えばガウス過程回帰などの手法により、微分可能な誤差予測モデルを生成する（ステップS15）。そして、データ生成部25は、誤差予測モデルを用いて、誤差が大きくなると予測されるラベルなし用例を生成し、出力部26は生成された用例を出力する（ステップS16）。そして、処理は終了する。

[0046] [変形例]

次に、第1実施形態の変形例について説明する。以下の変形例は、適宜組み合わせて第1実施形態に適用することができる。

(変形例1)

上記の実施形態では、ラベル生成部22は、予め用意された訓練済みの教師モデルを用いて、入力部21に入力されたラベルなし訓練用例にラベルを付与している。その代わりに、入力部21に対してラベル付き訓練用例が入

力される場合には、ラベル生成部 22 は、まず、ラベル付き訓練用例を用いて教師モデルを生成してもよい。また、ラベル生成部 22 では、教師モデルを用いてラベルを付与する代わりに、人手によりラベルを付与してもよい。また、上記の実施形態では、予測誤差算出部 23 は、ラベル生成部 22 がラベルを付与した訓練用例を用いて生徒モデルを生成しているが、その代わりに、予め用意された訓練済み生徒モデルを取得してもよい。

[0047] (変形例 2)

上記の実施形態では、出力部 26 は、誤差が大きくなると予測されるラベルなし用例を出力するが、出力部 26 の後段にラベル付与部を設けてもよい。こうすると、出力部 26 が出力したラベルなし用例に対して、ラベル付与部がラベルを付与することにより、生徒モデルの再訓練に使用可能なラベル付き訓練用例を生成することができる。なお、この場合、ラベル付与部は、ラベル生成部 22 で使用する教師モデルを用いてラベルを付与してもよく、ラベル生成部 22 が使用する教師モデルとは別の教師モデルを用いてラベルを付与してもよく、人手などによりラベルを付与してもよい。

[0048] <第 2 実施形態>

図 10 は、第 2 実施形態に係る情報処理装置 50 の機能構成を示すブロック図である。情報処理装置 50 は、入力手段 51 と、ラベル生成手段 52 と、誤差算出手段 53 と、誤差予測モデル生成手段 54 と、出力手段 55 とを備える。入力手段 51 は、特徴量からなる訓練用例を受け付ける。ラベル生成手段 52 は、訓練用例にラベルを付与する。誤差算出手段 53 は、ラベルが付与された訓練用例を用いて 1 つ以上の生徒モデルを生成し、当該生徒モデルによる予測とラベルとの誤差を算出する。誤差予測モデル生成手段 54 は、誤差を予測するモデルである誤差予測モデルを生成する。出力手段 55 は、誤差予測モデルに基づき、誤差が大きくなると予測される用例を出力する。

[0049] 図 11 は、第 2 実施形態の情報処理装置 50 による処理のフローチャートである。入力手段 51 は、特徴量からなる訓練用例を受け付ける（ステップ

S 2 1)。ラベル生成手段 5 2 は、訓練用例にラベルを付与する（ステップ S 2 2）。誤差算出手段 5 3 は、ラベルが付与された訓練用例を用いて 1 つ以上の生徒モデルを生成し、当該生徒モデルによる予測とラベルとの誤差を算出する（ステップ S 2 3）。誤差予測モデル生成手段 5 4 は、誤差を予測するモデルである誤差予測モデルを生成する（ステップ S 2 4）。出力手段 5 5 は、誤差予測モデルに基づき、誤差が大きくなると予測される用例を出力する（ステップ S 2 5）。

[0050] 第 2 実施形態の情報処理装置 5 0 によれば、訓練用例を用いて誤差予測モデルを生成し、誤差予測モデルに基づいて、誤差が大きくなると予測される用例が出力される。よって、出力された用例を用いて生徒モデルの再訓練を行うことにより、生徒モデルの精度を効率的に向上させることができる。

[0051] 上記の実施形態の一部又は全部は、以下の付記のようにも記載されうるが、以下には限られない。

[0052] (付記 1)

特徴量からなる訓練用例を受け付ける入力手段と、
前記訓練用例にラベルを付与するラベル生成手段と、
ラベルが付与された前記訓練用例を用いて 1 つ以上の生徒モデルを生成し、当該生徒モデルによる予測と前記ラベルとの誤差を算出する誤差算出手段と、
前記誤差を予測するモデルである誤差予測モデルを生成する誤差予測モデル生成手段と、
前記誤差予測モデルに基づき、誤差が大きくなると予測される用例を出力する出力手段と、
を備える情報処理装置。

[0053] (付記 2)

前記誤差予測モデル生成手段は、複数の前記訓練用例についての前記生徒モデルの予測とラベルとの誤差に基づいて、微分可能な誤差予測モデルを生成する付記 1 に記載の情報処理装置。

[0054] (付記3)

前記誤差予測モデルは回帰モデルであり、
前記出力手段は、前記回帰モデルにおける勾配に基づいて前記用例の誤差を予測する付記1又は2に記載の情報処理装置。

[0055] (付記4)

前記誤差予測モデルは、前記誤差の微分可能な平均及び分散を出力するモデルであり、
前記出力手段は、前記平均及び分散の少なくとも一方に基づいて、前記誤差が大きくなると予測される用例を出力する付記1又は2に記載の情報処理装置。

[0056] (付記5)

前記ラベル生成手段は、前記訓練用例を用いて生成された教師モデルを用いて、前記訓練用例にラベルを付与し、
前記誤差算出手段は、前記教師モデルによる予測である前記ラベルと、前記生徒モデルによる予測との誤差を算出する付記1乃至4のいずれか一項に記載の情報処理装置。

[0057] (付記6)

前記誤差算出手段は、前記訓練用例の少なくとも一部の用例を用いて前記1つ以上の生徒モデルを生成し、前記生徒モデルの生成に用いた用例とは異なる用例を用いて前記誤差を算出する付記1乃至5のいずれか一項に記載の情報処理装置。

[0058] (付記7)

前記誤差算出手段は、前記訓練用例から重複ありランダムサンプリングにより複数のサンプル群を生成し、前記サンプル群の各々を用いて生徒モデルを生成し、前記生徒モデル毎に、前記訓練用例に含まれるが前記サンプル群に含まれないデータについて前記誤差を算出し、前記複数の生徒モデルについて算出された誤差の平均を算出する付記1乃至5のいずれか一項に記載の情報処理装置。

[0059] (付記 8)

特徴量からなる訓練用例を受け付け、
前記訓練用例にラベルを付与し、
ラベルが付与された前記訓練用例を用いて1つ以上の生徒モデルを生成し、
当該生徒モデルによる予測と前記ラベルとの誤差を算出し、
前記誤差を予測するモデルである誤差予測モデルを生成し、
前記誤差予測モデルに基づき、誤差が大きくなると予測される用例を出力する情報処理方法。

[0060] (付記 9)

特徴量からなる訓練用例を受け付け、
前記訓練用例にラベルを付与し、
ラベルが付与された前記訓練用例を用いて1つ以上の生徒モデルを生成し、
当該生徒モデルによる予測と前記ラベルとの誤差を算出し、
前記誤差を予測するモデルである誤差予測モデルを生成し、
前記誤差予測モデルに基づき、誤差が大きくなると予測される用例を出力する処理をコンピュータに実行させるプログラムを記録した記録媒体。

[0061] 以上、実施形態及び実施例を参照して本発明を説明したが、本発明は上記実施形態及び実施例に限定されるものではない。本発明の構成や詳細には、本発明のスコープ内で当業者が理解し得る様々な変更をすることができる。

符号の説明

- [0062] 1 1 入力IF
1 2 プロセッサ
1 3 メモリ
1 4 記録媒体
1 5 データベース
2 1 入力部
2 2 ラベル生成部
2 3 予測誤差算出部

2 4 誤差予測モデル生成部

2 5 データ生成部

2 6 出力部

1 0 0 情報処理装置

請求の範囲

- [請求項1] 特徴量からなる訓練用例を受け付ける入力手段と、
前記訓練用例にラベルを付与するラベル生成手段と、
ラベルが付与された前記訓練用例を用いて1つ以上の生徒モデルを生成し、当該生徒モデルによる予測と前記ラベルとの誤差を算出する誤差算出手段と、
前記誤差を予測するモデルである誤差予測モデルを生成する誤差予測モデル生成手段と、
前記誤差予測モデルに基づき、誤差が大きくなると予測される用例を出力する出力手段と、
を備える情報処理装置。
- [請求項2] 前記誤差予測モデル生成手段は、複数の前記訓練用例についての前記生徒モデルの予測とラベルとの誤差に基づいて、微分可能な誤差予測モデルを生成する請求項1に記載の情報処理装置。
- [請求項3] 前記誤差予測モデルは回帰モデルであり、
前記出力手段は、前記回帰モデルにおける勾配に基づいて前記用例の誤差を予測する請求項1又は2に記載の情報処理装置。
- [請求項4] 前記誤差予測モデルは、前記誤差の微分可能な平均及び分散を出力するモデルであり、
前記出力手段は、前記平均及び分散の少なくとも一方に基づいて、前記誤差が大きくなると予測される用例を出力する請求項1又は2に記載の情報処理装置。
- [請求項5] 前記ラベル生成手段は、前記訓練用例を用いて生成された教師モデルを用いて、前記訓練用例にラベルを付与し、
前記誤差算出手段は、前記教師モデルによる予測である前記ラベルと、前記生徒モデルによる予測との誤差を算出する請求項1乃至4のいずれか一項に記載の情報処理装置。
- [請求項6] 前記誤差算出手段は、前記訓練用例の少なくとも一部の用例を用い

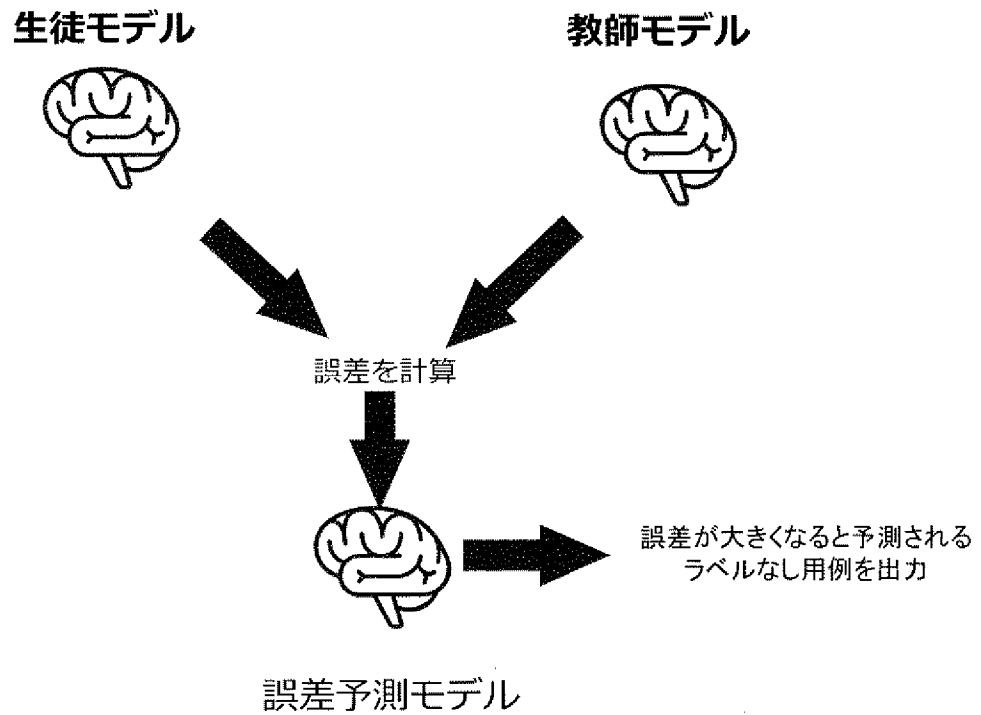
て前記1つ以上の生徒モデルを生成し、前記生徒モデルの生成に用いた用例とは異なる用例を用いて前記誤差を算出する請求項1乃至5のいずれか一項に記載の情報処理装置。

[請求項7] 前記誤差算出手段は、前記訓練用例から重複ありランダムサンプリングにより複数のサンプル群を生成し、前記サンプル群の各々を用いて生徒モデルを生成し、前記生徒モデル毎に、前記訓練用例に含まれるが前記サンプル群に含まれないデータについて前記誤差を算出し、前記複数の生徒モデルについて算出された誤差の平均を算出する請求項1乃至5のいずれか一項に記載の情報処理装置。

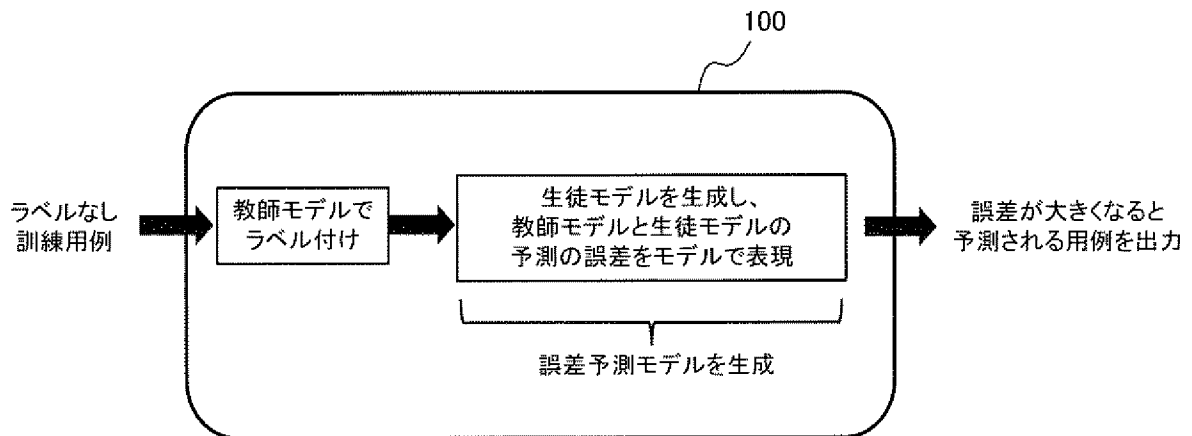
[請求項8] 特徴量からなる訓練用例を受け付け、
前記訓練用例にラベルを付与し、
ラベルが付与された前記訓練用例を用いて1つ以上の生徒モデルを生成し、当該生徒モデルによる予測と前記ラベルとの誤差を算出し、
前記誤差を予測するモデルである誤差予測モデルを生成し、
前記誤差予測モデルに基づき、誤差が大きくなると予測される用例を出力する情報処理方法。

[請求項9] 特徴量からなる訓練用例を受け付け、
前記訓練用例にラベルを付与し、
ラベルが付与された前記訓練用例を用いて1つ以上の生徒モデルを生成し、当該生徒モデルによる予測と前記ラベルとの誤差を算出し、
前記誤差を予測するモデルである誤差予測モデルを生成し、
前記誤差予測モデルに基づき、誤差が大きくなると予測される用例を出力する処理をコンピュータに実行させるプログラムを記録した記録媒体。

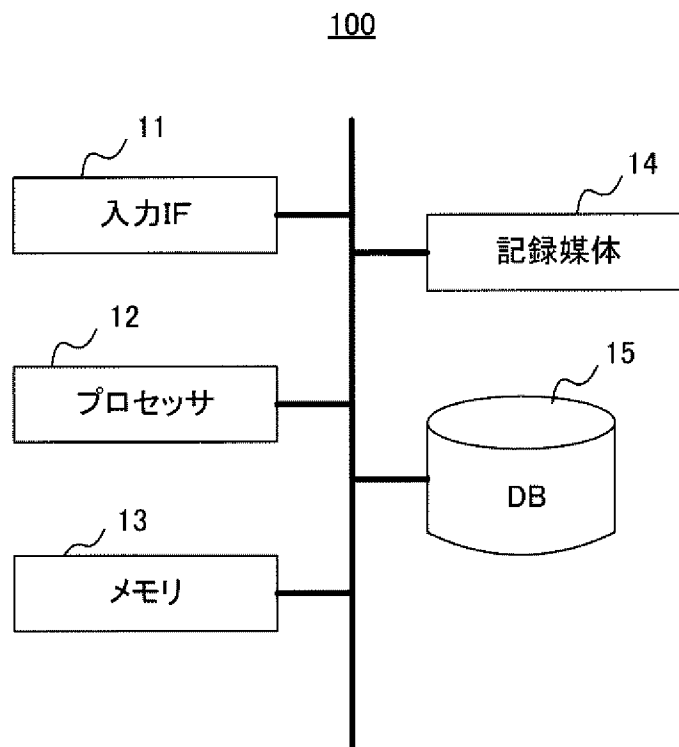
[図1]



[図2]

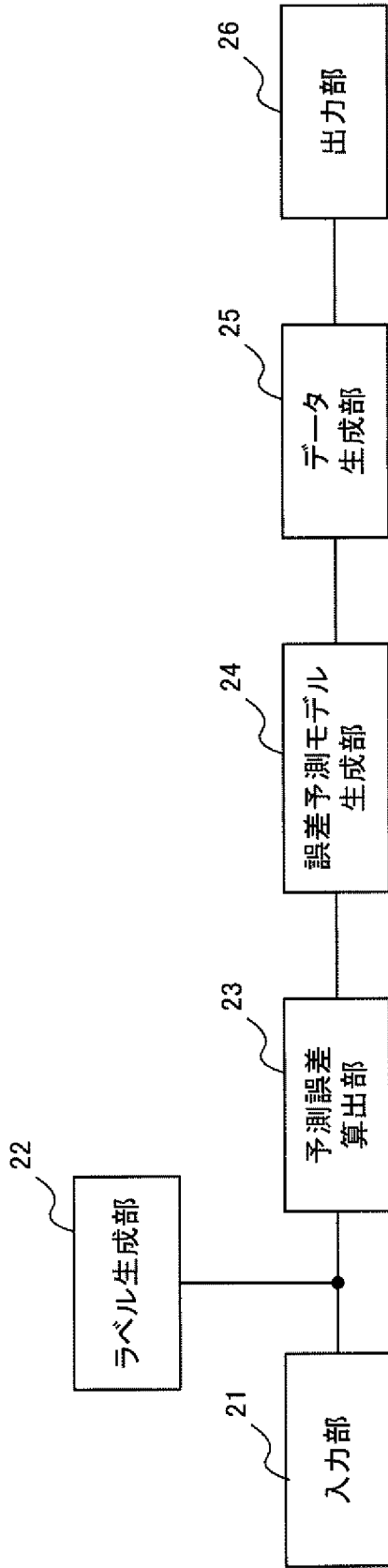


[図3]

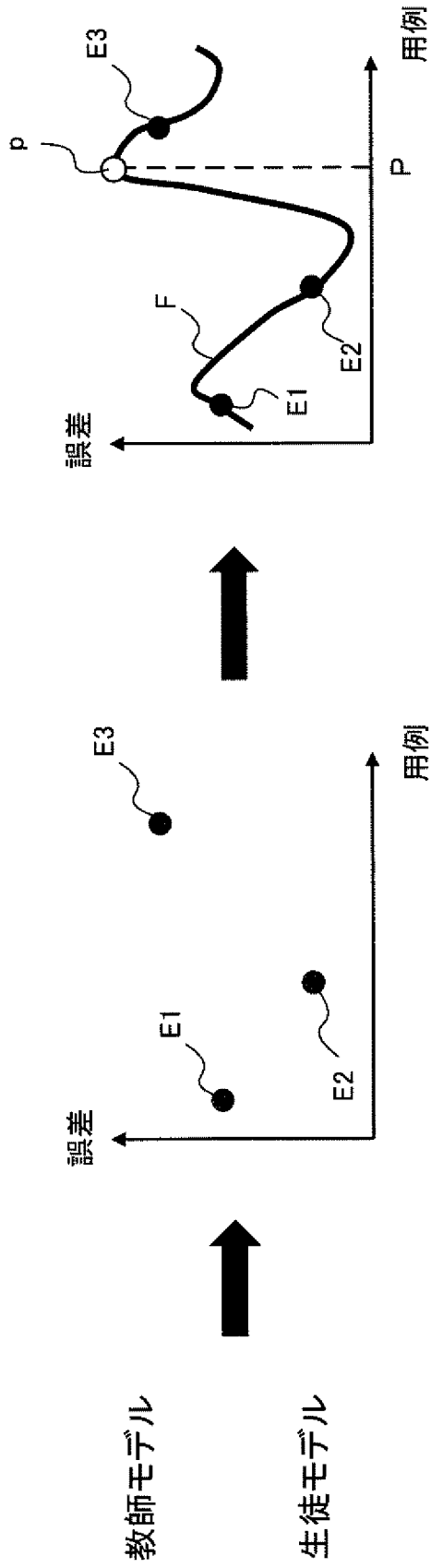


[図4]

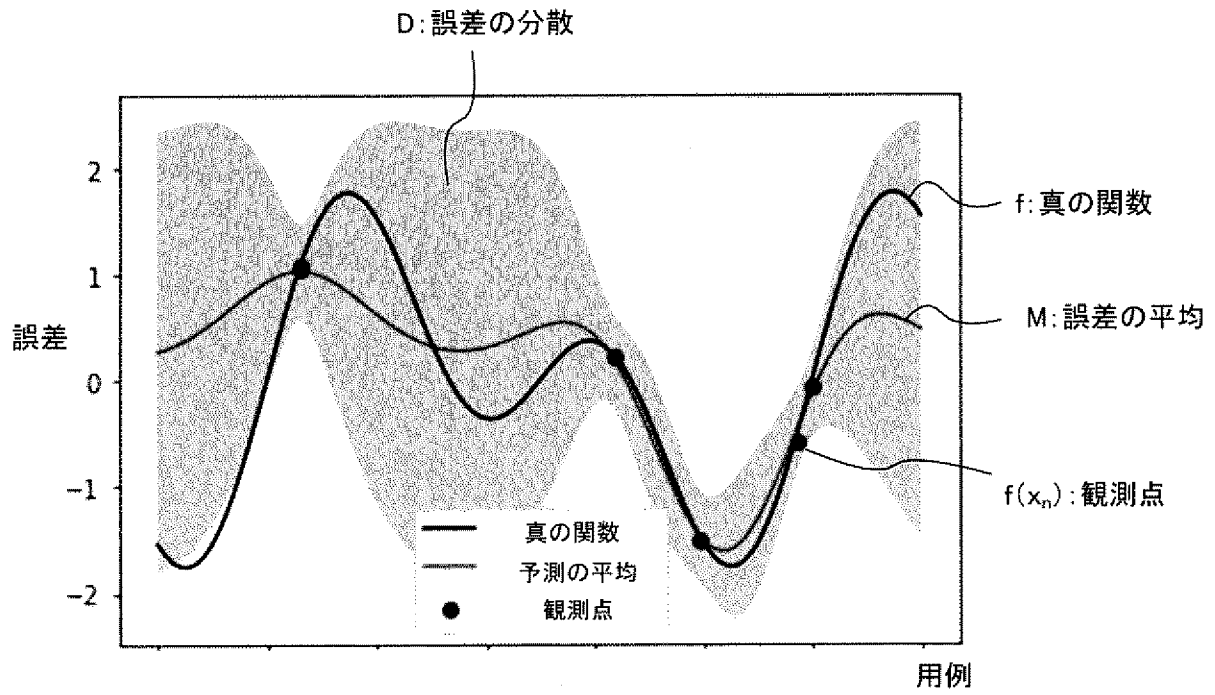
100



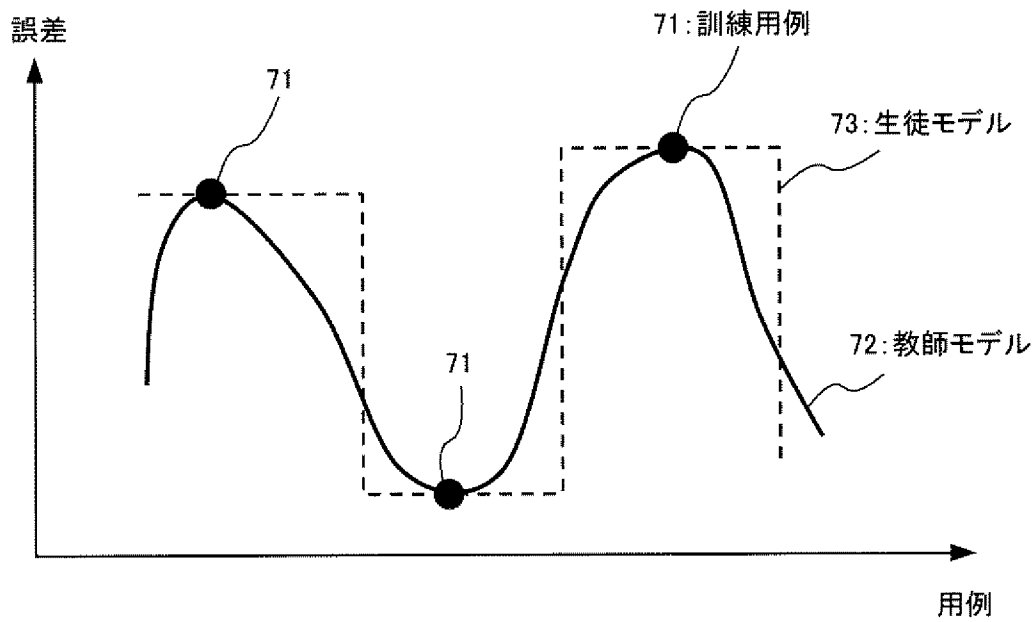
[図5]



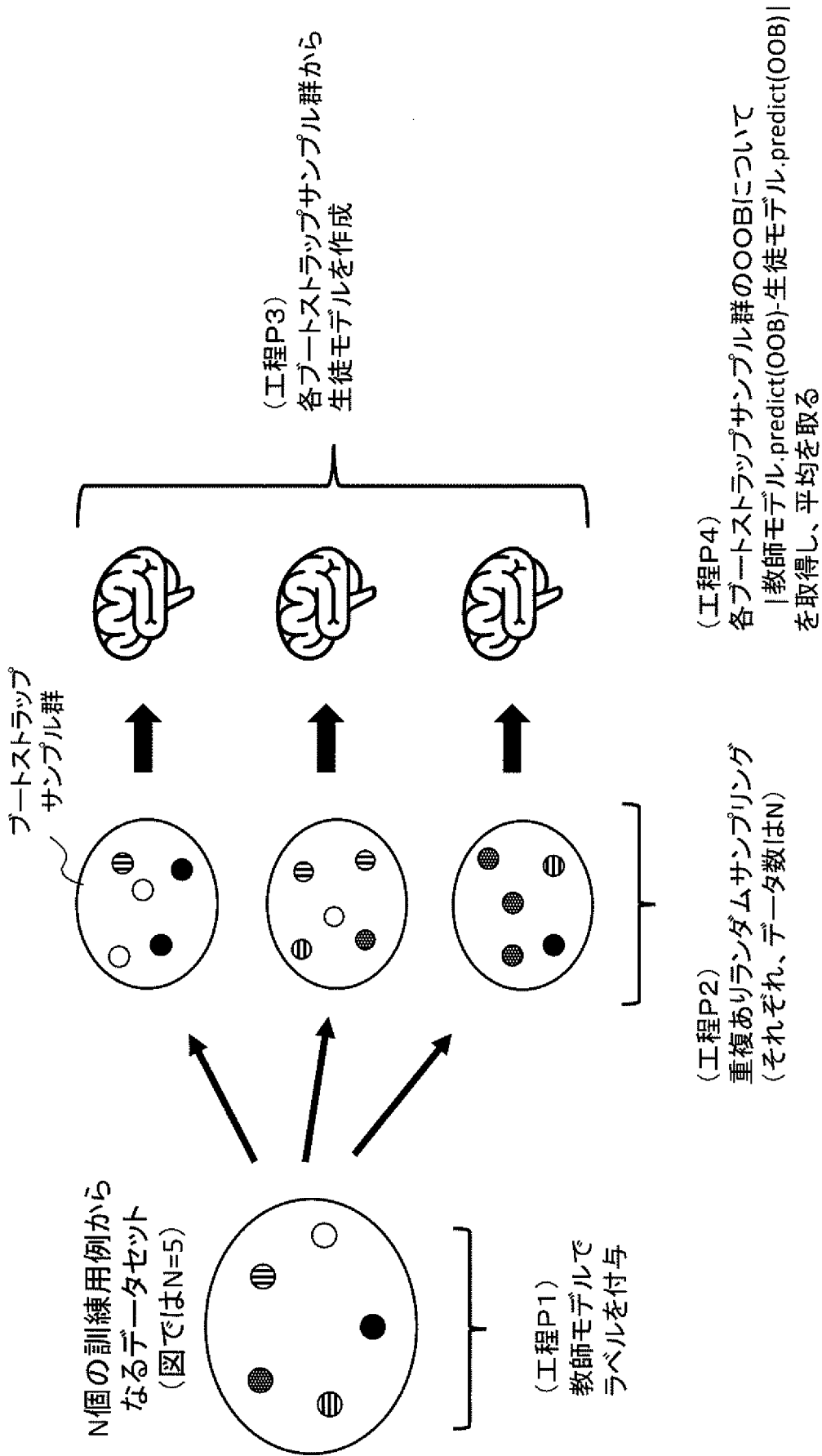
[図6]



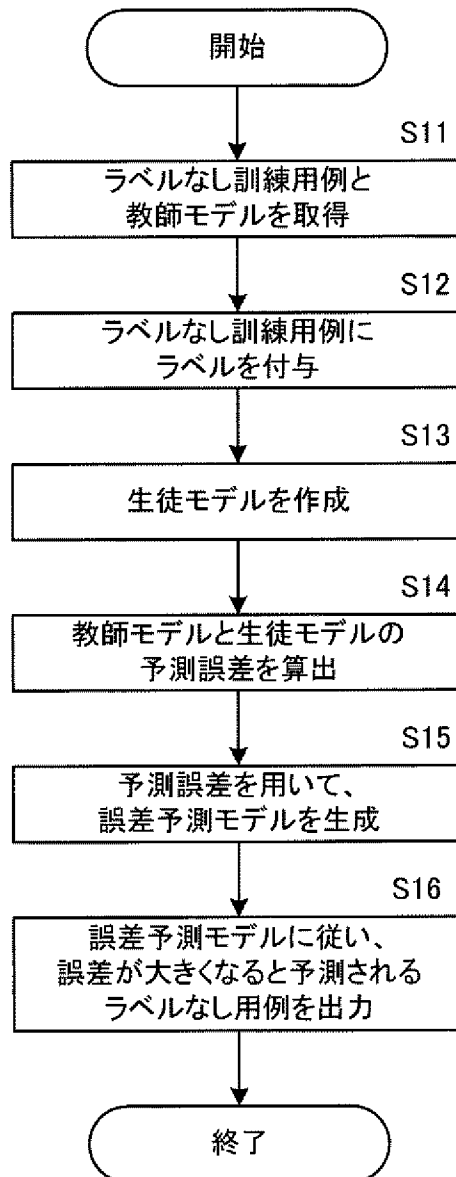
[図7]



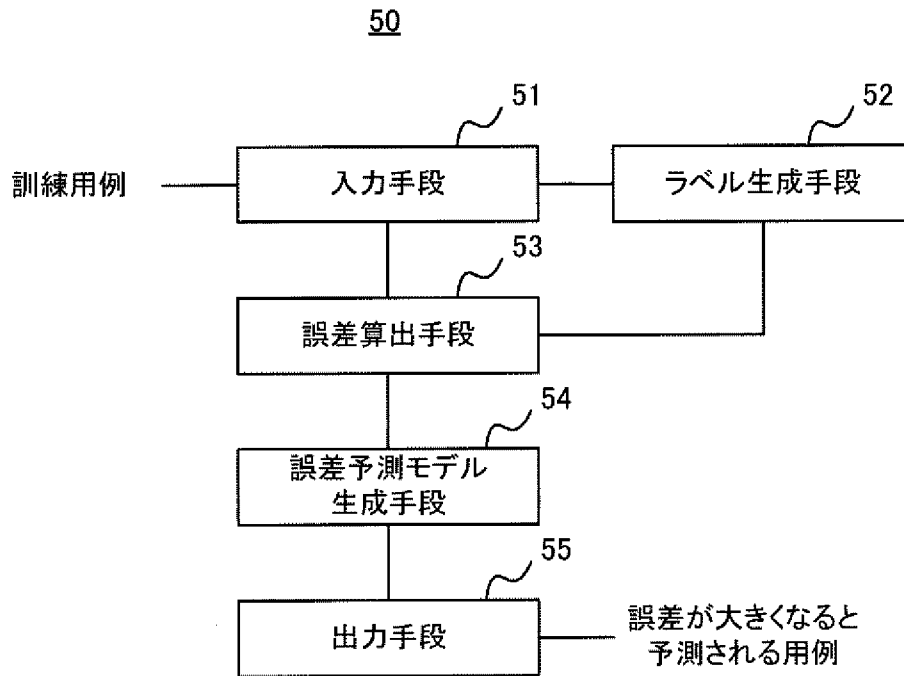
[図8]



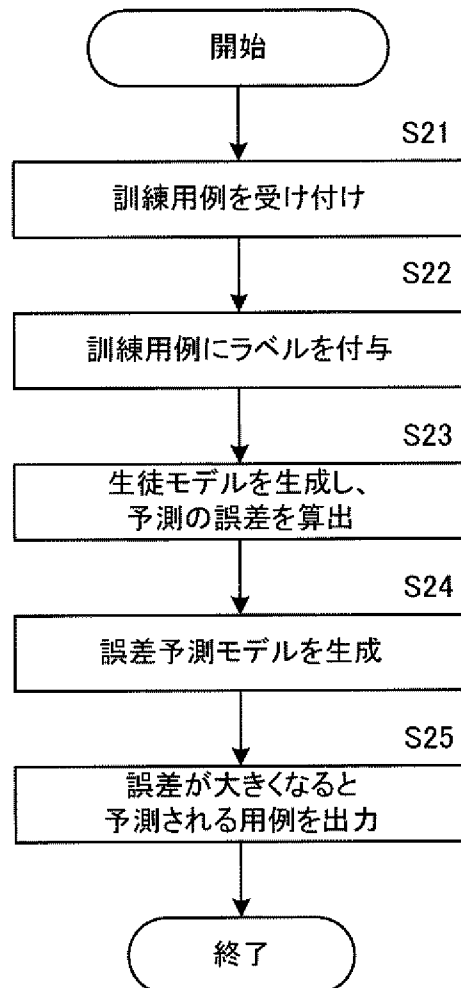
[図9]



[図10]



[図11]



INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2020/044483

<p>A. CLASSIFICATION OF SUBJECT MATTER Int.Cl. G06N20/00 (2019.01) i FI: G06N20/00130</p> <p>According to International Patent Classification (IPC) or to both national classification and IPC</p>																	
<p>B. FIELDS SEARCHED</p> <p>Minimum documentation searched (classification system followed by classification symbols) Int.Cl. G06N20/00</p> <p>Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched</p> <table style="width:100%; border:none;"> <tr> <td style="padding-left: 20px;">Published examined utility model applications of Japan</td> <td style="text-align: right;">1922-1996</td> </tr> <tr> <td style="padding-left: 20px;">Published unexamined utility model applications of Japan</td> <td style="text-align: right;">1971-2021</td> </tr> <tr> <td style="padding-left: 20px;">Registered utility model specifications of Japan</td> <td style="text-align: right;">1996-2021</td> </tr> <tr> <td style="padding-left: 20px;">Published registered utility model applications of Japan</td> <td style="text-align: right;">1994-2021</td> </tr> </table> <p>Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)</p>			Published examined utility model applications of Japan	1922-1996	Published unexamined utility model applications of Japan	1971-2021	Registered utility model specifications of Japan	1996-2021	Published registered utility model applications of Japan	1994-2021							
Published examined utility model applications of Japan	1922-1996																
Published unexamined utility model applications of Japan	1971-2021																
Registered utility model specifications of Japan	1996-2021																
Published registered utility model applications of Japan	1994-2021																
<p>C. DOCUMENTS CONSIDERED TO BE RELEVANT</p> <table border="1" style="width:100%; border-collapse: collapse;"> <thead> <tr> <th style="width:10%;">Category*</th> <th style="width:70%;">Citation of document, with indication, where appropriate, of the relevant passages</th> <th style="width:20%;">Relevant to claim No.</th> </tr> </thead> <tbody> <tr> <td align="center">Y</td> <td>YOO, D. et al., Learning loss for active learning., In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2019, pp. 93-102, particularly, 3. Method</td> <td align="center">1-9</td> </tr> <tr> <td align="center">Y</td> <td>JP 2019-159654 A (NATIONAL INSTITUTE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY) 19 September 2019 (2019-09-19), paragraph [0006]</td> <td align="center">1-9</td> </tr> <tr> <td align="center">Y</td> <td>WO 2020/162048 A1 (UNIVERSITY OF YAMANASHI) 13 August 2020 (2020-08-13), paragraph [0040]</td> <td align="center">1-9</td> </tr> <tr> <td align="center">Y</td> <td>JP 2016-528549 A (ASML NETHERLANDS BV) 15 September 2016 (2016-09-15), abstract</td> <td align="center">3</td> </tr> </tbody> </table>			Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.	Y	YOO, D. et al., Learning loss for active learning., In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2019, pp. 93-102, particularly, 3. Method	1-9	Y	JP 2019-159654 A (NATIONAL INSTITUTE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY) 19 September 2019 (2019-09-19), paragraph [0006]	1-9	Y	WO 2020/162048 A1 (UNIVERSITY OF YAMANASHI) 13 August 2020 (2020-08-13), paragraph [0040]	1-9	Y	JP 2016-528549 A (ASML NETHERLANDS BV) 15 September 2016 (2016-09-15), abstract	3
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.															
Y	YOO, D. et al., Learning loss for active learning., In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2019, pp. 93-102, particularly, 3. Method	1-9															
Y	JP 2019-159654 A (NATIONAL INSTITUTE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY) 19 September 2019 (2019-09-19), paragraph [0006]	1-9															
Y	WO 2020/162048 A1 (UNIVERSITY OF YAMANASHI) 13 August 2020 (2020-08-13), paragraph [0040]	1-9															
Y	JP 2016-528549 A (ASML NETHERLANDS BV) 15 September 2016 (2016-09-15), abstract	3															
<p><input checked="" type="checkbox"/> Further documents are listed in the continuation of Box C. <input checked="" type="checkbox"/> See patent family annex.</p>																	
<p>* Special categories of cited documents:</p> <table style="width:100%; border:none;"> <tr> <td style="width:50%; vertical-align: top;"> <p>“A” document defining the general state of the art which is not considered to be of particular relevance</p> <p>“E” earlier application or patent but published on or after the international filing date</p> <p>“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)</p> <p>“O” document referring to an oral disclosure, use, exhibition or other means</p> <p>“P” document published prior to the international filing date but later than the priority date claimed</p> </td> <td style="width:50%; vertical-align: top;"> <p>“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</p> <p>“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</p> <p>“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</p> <p>“&” document member of the same patent family</p> </td> </tr> </table>			<p>“A” document defining the general state of the art which is not considered to be of particular relevance</p> <p>“E” earlier application or patent but published on or after the international filing date</p> <p>“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)</p> <p>“O” document referring to an oral disclosure, use, exhibition or other means</p> <p>“P” document published prior to the international filing date but later than the priority date claimed</p>	<p>“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</p> <p>“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</p> <p>“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</p> <p>“&” document member of the same patent family</p>													
<p>“A” document defining the general state of the art which is not considered to be of particular relevance</p> <p>“E” earlier application or patent but published on or after the international filing date</p> <p>“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)</p> <p>“O” document referring to an oral disclosure, use, exhibition or other means</p> <p>“P” document published prior to the international filing date but later than the priority date claimed</p>	<p>“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</p> <p>“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</p> <p>“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</p> <p>“&” document member of the same patent family</p>																
<p>Date of the actual completion of the international search 21 January 2021</p>		<p>Date of mailing of the international search report 02 February 2021</p>															
<p>Name and mailing address of the ISA/ Japan Patent Office 3-4-3, Kasumigaseki, Chiyoda-ku, Tokyo 100-8915, Japan</p>		<p>Authorized officer</p> <p>Telephone No.</p>															

INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2020/044483

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	<p>稲津 佑, ほか, ガウス過程の導関数に基づく極小点の同定のための能動学習, 電子情報通信学会技術研究報告, 05 November 2018, vol. 118, no. 284, pp. 373-380, particularly, 2. 問題設定とガウス過程に関連する話題, (INATSU, Yu et al., Active learning for identifying local minimum points based on the derivative of Gaussian process, IEICE Technical Report), non-official translation (2. Topics related to problem setting and Gaussian process)</p>	4
Y	<p>ISHIGAMI, S., RandomForestとR package [online], slideshare, 01 February 2018, p. 34, [retrieved on 21 January 2021], Internet: <URL:https://www.slideshare.net/ShumaIshigami/randomforestr-package/34>, particularly, Out-Of-Bag(OOB)エラー, non-official translation (Random Forest and R package, Out-Of-Bag (OOB) error)</p>	7

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/JP2020/044483

JP 2019-159654 A	19 September 2019	(Family: none)
WO 2020/162048 A1	13 August 2020	(Family: none)
JP 2016-528549 A	15 September 2016	US 2016/0161864 A1 abstract WO 2015/018625 A1 CN 105452962 A

A. 発明の属する分野の分類（国際特許分類（IPC）） G06N 20/00(2019.01)i FI: G06N20/00 130		
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）） G06N20/00 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 - 2021年 日本国実用新案登録公報 1996 - 2021年 日本国登録実用新案公報 1994 - 2021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
Y	Y00, D., et al., Learning loss for active learning., In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2019, p. 93-102 特に、3. Method	1-9
Y	JP 2019-159654 A（国立研究開発法人情報通信研究機構）19.09.2019（2019 - 09 - 19） 段落[0006]	1-9
Y	WO 2020/162048 A1（国立大学法人山梨大学）13.08.2020（2020 - 08 - 13） 段落[0040]	1-9
Y	JP 2016-528549 A（エーエスエムエル ネザーランズ ビー. ブイ. ）15.09.2016（2016 - 09 - 15） 要約	3
Y	稲津 佑, ほか, ガウス過程の導関数に基づく極小点の同定のための能動学習, 電子情報通信学会技術研究報告, 2018.11.05, Vol.118, No.284, p.373-380 特に、2. 問題設定とガウス過程に関連する話題	4
<input checked="" type="checkbox"/> C欄の続きにも文献が列挙されている。 <input checked="" type="checkbox"/> 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 “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	21.01.2021	国際調査報告の発送日 02.02.2021
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） 石川 雄太郎 5B 5090 電話番号 03-3581-1101 内線 3545	

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
Y	ISHIGAMI, S., RandomForestとR package [online], slideshare, 2018.02.01, p.34, [検索日 2021.01.21], Internet: <URL: https://www.slideshare.net/ ShumaIshigami/randomforestr-package/34 > 特に、Out-of-Bag(OOB)エラー	7

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/JP2020/044483

引用文献	公表日	パテントファミリー文献	公表日
JP 2019-159654 A	19.09.2019	(ファミリーなし)	
WO 2020/162048 A1	13.08.2020	(ファミリーなし)	
JP 2016-528549 A	15.09.2016	US 2016/0161864 A1 ABSTRACT	
		WO 2015/018625 A1	
		CN 105452962 A	