

令和 6 年 6 月 5 日現在

機関番号：32689

研究種目：基盤研究(C)（一般）

研究期間：2019～2023

課題番号：19K04914

研究課題名（和文）ビジネス価値創造のためのデータ解析プラットフォームと時変協調フィルタリングの研究

研究課題名（英文）Research on Data Analysis Platform for Business Value Creation and Time-Varying Collaborative Filtering

研究代表者

小林 学（Kobayashi, Manabu）

早稲田大学・データ科学センター・教授

研究者番号：80308204

交付決定額（研究期間全体）：（直接経費） 3,400,000円

研究成果の概要（和文）：本研究では企業の持つクローズドデータを、外部に出力することなくサーバ上でセキュアな環境のもとで解析を実施するプラットフォーム(DAPF)について検討並びに運用を行なった。クラウドDAPFでは実データの解析を行う延べ5つのプロジェクトによる実運用を実施し、またオンプレミスDAPFではコンテナによるマルチ分析環境や学外接続など拡張性の検証を行なった。また協調フィルタリングに関して、顧客や商品の属性情報と時間変化する情報、消費行動など異なる種類のデータに複数の潜在変数を仮定し、それらの変数間の関係によりこれらのデータを統合的に扱う潜在構造モデルの提案を行なった。また実験によりその有効性を示した。

研究成果の学術的意義や社会的意義

企業の持つビッグデータを大学の研究者が安全に分析を行うプラットフォームは、企業にとっては新たな価値創造の機会として、大学の研究者にとっては学術研究の発展のために大いに役立つ。この時外部へデータを漏洩させない仕組みが非常に重要である。また個人（ユーザ）が商品（サービス）に対して行う日々の消費行動について、それぞれに対する潜在変数と観測変数の関係としてある種の統合的な視点でモデル化及び推定を行う方法を示した本研究は、学術的および社会的な観点の両者に対して意義がある。

研究成果の概要（英文）：In this research, we constructed a platform (DAPF) that enables the analysis of proprietary big data held by companies in a secure environment on a server without exporting the data outside the server. The cloud DAPF effectively conducted practical operations in a total of five projects that analyzed real data. Furthermore, the on-premise DAPF verified scalability, such as multiple analysis environments using containers and external connections via VPN. Regarding collaborative filtering, we proposed a latent structure model that assumes multiple latent variables for various types of data, including customer and product attribute information, time information, and consumption behavior. This model integrates these data by considering the relationships between the observed and latent variables. We also evaluated the effectiveness of this model.

研究分野：統計的学習理論

キーワード：データ解析プラットフォーム 潜在構造分析 協調フィルタリング ビッグデータ 統計的機械学習

様式 C-19、F-19-1 (共通)

1. 研究開始当初の背景

(1) 多種多様かつ大量なデータを扱えるようになり、データサイエンスの重要性が高まる中、ビッグデータを持つ企業はこれを有効に活用して新たなビジネス価値創造を目指す。十分にデータ分析が可能な人材を常に確保することが難しい状況であった。ここで一般にデータは広く利用可能なオープンデータと、主に企業が独自に所有するクローズドデータに分かれる。大学の研究者にとって企業の持つクローズドデータの学術的な価値が高いケースも多い。このような背景から企業の持つデータを大学研究者に活用してもらった産学連携の動きがありつつも、特にセキュリティ面は大きな課題であった。

(2) 実際の企業データには顧客の購買履歴、価格などを含む商品情報、WEB アクセス回数、等々の様々なデータが時点とともに蓄えられている。このとき協調フィルタリングの分野においては、顧客と商品の消費行動の直接的な数理モデル化を行う研究が種々存在するが、顧客情報、商品情報、時点を含む日々の消費行動履歴などのデータを統一的な観点から扱う方法は明らかではなかった。

(3) データ分析を行うための数理モデルとして、高い説明性を持ちつつ十分な精度を実現する研究が着目されつつあった。またデータ分析の適用範囲が急速に拡大するに従い、精度を高く保ったまま計算量を削減する方法が求められ始めていた。

2. 研究の目的

(1) 大学においてクローズドデータを解析対象とする場合にはあらかじめ企業側で個人情報やマスクされることが多いが、それでも万一データが（その一部といえども）外部に流出するようなことがあった場合には、その企業並びに大学に対する社会的イメージが大幅に下がる大きなリスクが存在する。このため、ビジネス価値を創造するためのデータを用いた産学連携に踏み切れない企業が少なくない。そのためにデータを外部に持ち出すことができないデータ解析プラットフォーム(DAPF)を用いることにより、大学教員や連携研究者、学生は通常利用することができないデータを用いて研究を行うことが可能となる。また DAPF に対して効果的な運用やシステムの拡張について検討し、明らかにする。

(2) 協調フィルタリングにおける顧客と商品それぞれの情報、さらに消費行動などの時点を含む種々の観測データに対して、高い説明性を持つ統合的な潜在構造モデルを明らかにする。またそのようなモデルに対して構造推定および予測について十分な精度を実現する推定アルゴリズムを構築する。

(3) 分類問題（目的変数が質的変数の場合）におけるクラス数が多い場合に対する予測を行う際に、予測精度をなるべく下げずに計算量を削減する。

3. 研究の方法

(1) DAPF に関して、クラウドサーバを利用してデータ解析を行う DAPF を CDAPF と呼ぶことにする。一方オンプレミスのデータ解析用サーバを置き、このサーバを用いた DAPF をオンプレミス DAPF (ODAPF) と呼ぶことにする。CDAPF は実際のプロジェクトにおける運用することに主眼を置き、セキュリティ要件の他効果的な運用方法の検討を行う。一方 ODAPF はオンプレミスの利点を活かし、自由な構成でその拡張性や利便性について検討を行う。

(2) 協調フィルタリングに対しては、顧客および商品にそれぞれ潜在変数を仮定する。さらに種々の消費行動がこれらの潜在変数との関係で表現可能な数理モデルの構築を行う。また時点により変化する観測データに対する潜在変数も考慮する。さらに提案した数理モデルに対する構造推定および予測を行うアルゴリズムの開発を行い、提案モデルの有効性の評価を行う。

(3) データに対する説明性を持ち、また実際に予測に対する十分な予測精度を持つ手法に一般化加法モデルがある。さらに目的変数が質的変数である分類問題において、分類クラスの数が多くなったときに精度を高める手法に誤り訂正出力符号 (Error Correcting Output Codes : ECOC と略す) がある。これらの従来手法に対して、十分な説明性の確保と精度を高めること、さらには計算量を低減する方法を明らかにする。また実験による評価および理論的な評価を行う。

4. 研究成果

(1) クラウド利用データ解析プラットフォーム(CDAPF)に関してその概要を図 1 に示す¹。システムの構成要素は大別して 1)「Cloud 分析環境」、2)「データアクセス室」、3)「データ分析 PC」がある。「Cloud 分析環境」にはデータ提供元の企業によってマスキング済みのデータがアップ

ロードされる。大学側では分析を実施する利用者（以下利用者）は、「データアクセス室」において分析専用の「データ分析 PC」を用いて「Cloud 分析環境」にアクセスし、データ分析を実施する。ここで「データ分析 PC」は「Cloud 分析環境」のみとしか接続できない設定としており、また「データ分析 PC」は管理者以外がファイルを取り出すことができない設定にしている。「Cloud 分析環境」は外部のインターネットには接続できない設定にしておき、セキュリティの確保を行う。また「Cloud 分析環境」では R や Python が利用可能としている。また「データ分析 PC」においては Excel の他、必要に応じて STATA 等の分析ソフトウェアを利用可能とした。

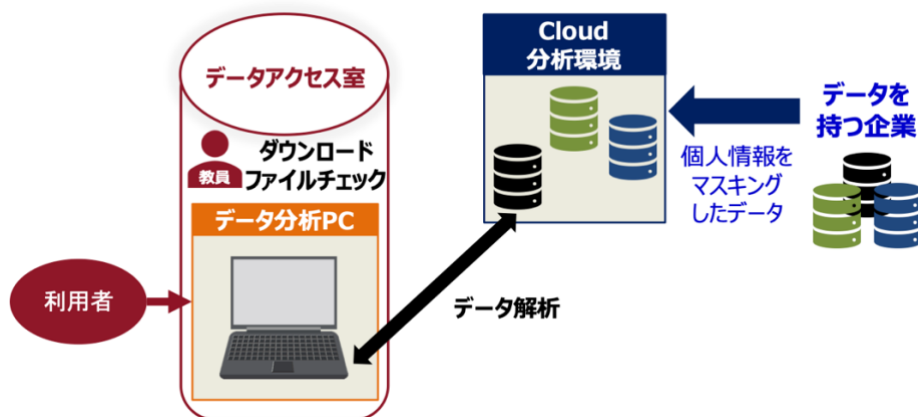


図 1：CDAPF のシステムの概要

CDAPF を用いて延べ 5 件のプロジェクトの実運用を実施し、利用者の分析効率をなるべく下げずに、データ流出を起こさないためのセキュリティについてその要件を明らかにした。具体的にはプロジェクト管理、利用者管理、データアクセス室入室管理、データ分析 PC 管理、Cloud 分析環境管理、データ分析 PC セキュリティ設定、結果ファイルダウンロードチェック、結果ファイル受渡し管理など多岐にわたる。なお、必要なデータをタイムリーに Cloud 分析環境に用意することが DAPF として重要であるが、これは企業の負担となる点も明らかとなった。このように企業の負担軽減を検討する必要がある点は今後の大きな課題である。

またオンプレミス利用データ解析プラットフォーム(ODAPF)に関して、その概要を図 2 に示す。ODAPF は拡張性を重視して、種々の利用者が異なるプロジェクトに参画する場面を想定し、さらに利用する解析環境も様々な種類が存在する場合を想定している。このときセキュリティのため、データ解析用サーバからの直接的なファイルダウンロードとインターネットアクセスを禁止するのは CDAPF と同じ考え方であるが、ファイアウォールを用いた VPN 接続を可能

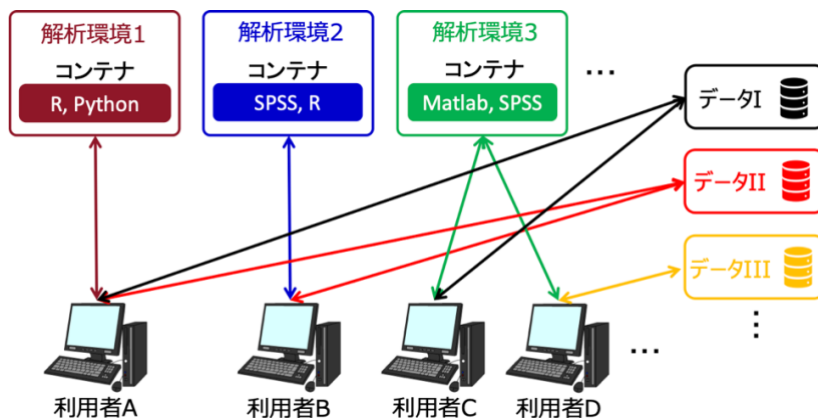


図 2：ODAPF のシステムの概要

とするなど、利用効率の向上も同時に目指した。これらは具体的には伊藤忠テクノソリューションズ(CTC)の CINAPS により実現されている。具体的には Docker のコンテナ技術により個別に異なる解析環境をコンテナとして用意し、利用者は自身の利用したいコンテナを選択する。またデータの利用については例えばユーザのグループによるアクセス権限およびパーミッションの設定による実現のほか、より汎用に利用者ごとにストレージ領域を仮想的に分離するマルチテナント化を実現する仮想ストレージリソースによる実現も可能である。結果的に自由な解析環境の構築を実現でき、また利用者やプロジェクトごとに別のデータを利用可能とすることにより、高いセキュリティを担保しつつ汎用性の高い DAPF を実現可能であることが分かった。ODAPF における課題を挙げるとするならば、新しいコンテナを用意するためにはある程度サーバに関する専門的な知見を必要とする点がある。また必要なデータの準備における企業負担については CDAPF と同様である。

(2) 時変の観測変数を含む協調フィルタリングに関して述べる。顧客はその属性情報など、顧客

ごとに観測できる変数（顧客観測変数）を持つ．同様に商品もジャンルやサイズ，機能などそれぞれ観測変数を持っており，これを商品観測変数と呼ぶ．このとき商品観測変数には，例えば商品価格や割引の有無やその額，購入季節や購入された曜日など時点により変化を伴う観測変数が含まれているものとする．次に顧客および商品の観測変数に対応する潜在変数を仮定し，それぞれ顧客潜在変数，商品潜在変数と呼ぶことにする．このとき，商品潜在変数には時点により変化する潜在変数と，商品によって変化しない潜在変数の 2 種類が存在するものとする．さらに別の観測変数として，顧客の商品に対する消費行動を表す観測変数が存在する．このとき，それぞれの観測変数は潜在変数の関数から確率的に生起するモデルとする．図 3 にこれらの関係を表現するベイジアンネットワークを示す．

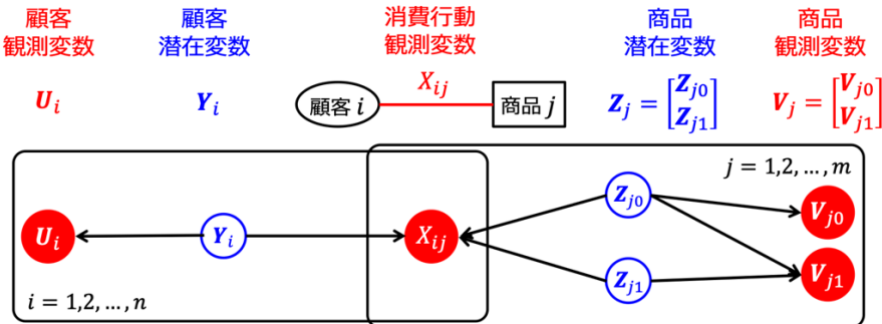


図 3：時変観測変数を含む潜在構造モデルのベイジアンネットワーク

ここで i と j はそれぞれ顧客と商品のインデックスを表し，顧客，商品および消費行動に対応する観測変数をそれぞれ U_i, V_j, X_{ij} としている．ただし V_j は時変でない観測変数 V_{j0} と時変の観測変数 V_{j1} に区別される．また顧客および商品に対応する潜在変数はそれぞれ Y_i, Z_j だが， Z_j は時変でない Z_{j0} と時変の Z_{j1} に分かれる点に注意されたい．ここで観測変数 U_i は Y_i から因子分析モデルとして生成されるものとする．一方観測変数 V_j は Z_j から構造方程式モデルとして生成されるものとする．さて X_{ij} は量的変数の場合と質的変数の場合が考えられる．量的変数の場合には X_{ij} は $Y_i^T Z_j$ に正規誤差が加わったモデル 1 を考え，質的変数の場合には $Y_i^T Z_j$ に対するロジスティックモデルにより X_{ij} の確率が定まるモデル 2 を考える．ここでモデル 1 は文献 3 のモデルと対応し，モデル 2 は文献 4 のモデルと対応しており，それぞれモンテカルロ EM アルゴリズムにより効果的にパラメータを推定可能であり，これらのモデルが人工データおよび実データに対して有効であることを明らかにした．本研究の潜在構造モデルは生成観測メカニズムを陽に仮定して統計的に構造推定および予測を行う手法であるため，得られたパラメータおよび推定した構造は分析者にとってその意味を十分に解釈および説明可能である．またそのため観測変数に欠測が含まれていても問題なく推定が可能である点は重要な利点である．

(3) (2)でも説明したように観測データの生成観測メカニズムを仮定して，また得られた構造推定の結果に対して分析者がそのパラメータの意味を十分に解釈および説明が可能であることは非常に重要である．そこで本研究では回帰モデルとして，図 4 のようなデータの生成観測メカニズムを仮定する．

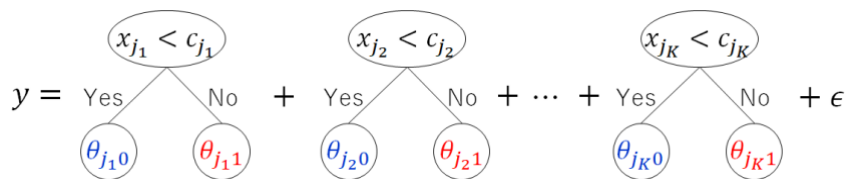


図 4：多数の回帰木から目的変数が決定される回帰モデル

ここで K は回帰木の数を表しており， j_1, \dots, j_K は説明変数のインデックスを表している． θ や c はモデルのパラメータである．

このような回帰モデルに対して，文献 5 ではまずデータが一つの回帰木 ($K = 1$ の場合) から統計的に生成されるモデルを仮定し，これに対するベイズリスクを最小とする最適な予測を示した．結果的にこの予測は数多くの回帰木の重みづけとなり，一般化加法モデルの一種となることを示した．さらに，この方法を予測の残差に対して繰り返し適用するブースティング手法の提案を行ない．結果的にこの手法も一般化加法モデルとなる．また従来の一般化加法モデルよりも良い結果となることを実験により示した．なおこの手法は分類問題や 2 次の交互作用項に対しても適用することができ，十分統計量を高速に計算する FAST アルゴリズムを用いることにより効果的に拡張可能である利点も併せ持つ．

(4) 目的変数が質的変数である分類問題において，分類のクラス数が二値より多い場合に利用される ECOC 法に対して，その構成法と実験の評価に関する研究 6，ある種の対称性を満たす符号を用いた場合の性能に関する理論的評価に関する研究 7, 8，符号をランダムに構成した場合の

理論的評価に関する研究⁹を行なった。特に文献 9 において、ランダムに構成した符号を用いた ECOC 法の平均的性能は、二値分類器の数が分類するクラス数の対数オーダー分だけ用意すると、いくつかの仮定と条件の下で漸近的には最適な事後確率推定を達成可能であることを示した。従来 ECOC 法では分類のクラス数とほぼ同じかこの多項式オーダー分だけ二値分類器を用いているため、性能をほぼ劣化させずに分類の計算量を劇的に低減できることを理論的に明らかにしている。これはこの分野の研究において非常に重要な結果である。一方で、ランダムに符号を構成した場合の平均的な性能について述べているため、具体的な符号の構成や復号法については、この結果を鑑みた上で継続して検討を行う必要がある。

<引用文献>

- ① 小林 学, 中原悠太, 佐藤絵里子, 高谷親信, 塚本好宏, “銀行データを活用するための産学連携研究プラットフォームの構築”, 情報処理学会第 84 回全国大会予稿集, March 2022
- ② データ分析基盤の実践構築セミナー. <https://www.sbbit.jp/eventinfo/59073/>
- ③ K. Mikawa, M. Kobayashi, M. Goto, S. Hirasawa, “Learning and Estimation of Latent Structural Models Based on Between-Data Metrics”, Proc. of 2022 IEEE International Conference on Systems, Man, and Cybernetics (SMC2022), Oct. 2022
- ④ K. Mikawa, M. Kobayashi, T. Sasaki, A. Manada, “A Study on Latent Structural Models for Binary Relational Data with Attribute Information”, Nonlinear Theory and Its Applications, IEICE E15-N(2), April 2024
- ⑤ 小林 学, 松嶋敏泰, 平澤茂一, “回帰木のベイズ的アプローチに基づく GAM/GA2M”, 日本経営工学会春季大会予稿集, May 2021
- ⑥ S. Hirasawa, G. Kumoi, H. Yagi, M. Kobayashi, M. Goto, H. Inazumi, “Construction Methods for Error Correcting Output Codes Using Constructive Coding and Their System Evaluations”, Proc. of 2022 IEEE International Conference on Systems, Man, and Cybernetics (SMC2022) Oct. 2022
- ⑦ 雲居 玄道, 八木 秀樹, 小林 学, 後藤 正幸, 平澤 茂一, “多値分類問題における ECOC 法の最適性に関する一考察”, 情報処理学会論文誌数理モデル化と応用 (TOM) 14(3) pp.1-10, 2021
- ⑧ G. Kumoi, H. Yagi, M. Kobayashi, M. Goto, S. Hirasawa, “Performance Evaluation of Error-Correcting Output Coding Based on Noisy and Noiseless Binary Classifiers”, International Journal of Neural Systems 33(02), Feb. 2023
- ⑨ 小林 学, 雲居 玄道, 八木 秀樹, 平澤 茂一, “ECOC 法に基づく多値分類の平均的性能について”, 第 46 回情報理論とその応用シンポジウム予稿集, pp.83-88, Nov. 2023

5. 主な発表論文等

〔雑誌論文〕 計26件（うち査読付論文 26件 / うち国際共著 0件 / うちオープンアクセス 6件）

1. 著者名 YASUDA Goki, SUKO Tota, KOBAYASHI Manabu, MATSUSHIMA Toshiyasu	4. 巻 E106.A
2. 論文標題 Asymptotic Evaluation of Classification in the Presence of Label Noise	5. 発行年 2023年
3. 雑誌名 IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences	6. 最初と最後の頁 422 ~ 430
掲載論文のDOI (デジタルオブジェクト識別子) 10.1587/transfun.2022TAP0013	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -
1. 著者名 Kumoi Gendo, Yagi Hideki, Kobayashi Manabu, Goto Masayuki, Hirasawa Shigeichi	4. 巻 33
2. 論文標題 Performance Evaluation of Error-Correcting Output Coding Based on Noisy and Noiseless Binary Classifiers	5. 発行年 2023年
3. 雑誌名 International Journal of Neural Systems	6. 最初と最後の頁 -
掲載論文のDOI (デジタルオブジェクト識別子) 10.1142/s0129065723500041	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -
1. 著者名 Hirasawa Shigeichi, Kumoi Gendo, Yagi Hideki, Kobayashi Manabu, Goto Masayuki, Inazumi Hiroshige	4. 巻 -
2. 論文標題 Construction Methods for Error Correcting Output Codes Using Constructive Coding and Their System Evaluations	5. 発行年 2022年
3. 雑誌名 Proc. of 2022 IEEE International Conference on Systems, Man, and Cybernetics	6. 最初と最後の頁 -
掲載論文のDOI (デジタルオブジェクト識別子) 10.1109/smc53654.2022.9945163	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -
1. 著者名 Kumoi Gendo, Yagi Hideki, Kobayashi Manabu, Hirasawa Shigeichi	4. 巻 -
2. 論文標題 Effect of Hamming Distance on Performance of ECOC with Estimated Binary Classifiers	5. 発行年 2022年
3. 雑誌名 Proc. of 2022 IEEE International Conference on Systems, Man, and Cybernetics	6. 最初と最後の頁 -
掲載論文のDOI (デジタルオブジェクト識別子) 10.1109/SMC53654.2022.9945575	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 Mikawa Kenta, Kobayashi Manabu, Goto Masayuki, Hirasawa Shigeichi	4. 巻 -
2. 論文標題 Learning and Estimation of Latent Structural Models Based on between-Data Metrics	5. 発行年 2022年
3. 雑誌名 Proc. of 2022 IEEE International Conference on Systems, Man, and Cybernetics	6. 最初と最後の頁 -
掲載論文のDOI (デジタルオブジェクト識別子) 10.1109/SMC53654.2022.9945506	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 Gendo Kumoi, Hideki Yagi, Manabu Kobayashi, Masayuki Goto, Shigeichi Hirasawa	4. 巻 -
2. 論文標題 Performance Evaluation of Ecoc Considering Estimated Probability of Binary Classifiers	5. 発行年 2022年
3. 雑誌名 Proceeding of WORLD CIST 2022	6. 最初と最後の頁 -
掲載論文のDOI (デジタルオブジェクト識別子) なし	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 Gendo Kumoi, Hideki Yagi, Manabu Kobayashi, Masayuki Goto, Shigeichi Hirasawa	4. 巻 -
2. 論文標題 Performance Evaluation of Ecoc Considering Estimated Probability of Binary Classifiers	5. 発行年 2022年
3. 雑誌名 Proceeding of WORLD CIST	6. 最初と最後の頁 -
掲載論文のDOI (デジタルオブジェクト識別子) なし	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 Katsuyuki Umezawa, Makoto Nakazawa, Manabu Kobayashi, Yutaka Ishii, Michiko Nakano, Shigeichi Hirasawa	4. 巻 -
2. 論文標題 Research Results on System Development of the Research Project of a Self-Study System for Language Learning	5. 発行年 2022年
3. 雑誌名 Proceeding of the The IEEE World Engineering Education Conference	6. 最初と最後の頁 -
掲載論文のDOI (デジタルオブジェクト識別子) なし	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 Katsuyuki Umezawa, Makoto Nakazawa, Manabu Kobayashi, Yutaka Ishii, Michiko Nakano, Shigeichi Hirasawa	4. 巻 -
2. 論文標題 Comparison Experiment of Learning State between Visual Programming Language and Text Programming Language	5. 発行年 2021年
3. 雑誌名 Proceeding of the IEEE International Conference on Teaching, Assessment and Learning for Engineering	6. 最初と最後の頁 -
掲載論文のDOI (デジタルオブジェクト識別子) なし	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 雲居 玄道, 八木 秀樹, 小林 学, 後藤 正幸, 平澤 茂一	4. 巻 14
2. 論文標題 多値分類問題におけるECOC法の最適性に関する一考察	5. 発行年 2021年
3. 雑誌名 情報処理学会論文誌数理モデル化と応用	6. 最初と最後の頁 1-10
掲載論文のDOI (デジタルオブジェクト識別子) なし	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 Nao Dobashi, Shota Saito, Yuta Nakahara, Toshiyasu Matsushima	4. 巻 23
2. 論文標題 Meta-Tree Random Forest: Probabilistic Data-Generative Model and Bayes Optimal Prediction	5. 発行年 2021年
3. 雑誌名 Entropy	6. 最初と最後の頁 768-768
掲載論文のDOI (デジタルオブジェクト識別子) 10.3390/e23060768	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 SUZUKI Kairi, KAMATSUKA Akira, MATSUSHIMA Toshiyasu	4. 巻 E103.A
2. 論文標題 A Bayesian Decision-Theoretic Change-Point Detection for i.p.i.d. Sources	5. 発行年 2020年
3. 雑誌名 IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences	6. 最初と最後の頁 1393 ~ 1402
掲載論文のDOI (デジタルオブジェクト識別子) 10.1587/transfun.2020TAP0009	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 NAKAHARA Yuta、MATSUSHIMA Toshiyasu	4. 巻 E103.A
2. 論文標題 Analysis of Decoding Error Probability of Spatially “Mt. Fuji” Coupled LDPC Codes in Waterfall Region of the BEC	5. 発行年 2020年
3. 雑誌名 IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences	6. 最初と最後の頁 1337 ~ 1346
掲載論文のDOI (デジタルオブジェクト識別子) 10.1587/transfun.2020TAP0010	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 Saito Shota、Matsushima Toshiyasu	4. 巻 -
2. 論文標題 Evaluation of Error Probability of Classification Based on the Analysis of the Bayes Code	5. 発行年 2020年
3. 雑誌名 Proc. of 2020 IEEE International Symposium on Information Theory	6. 最初と最後の頁 2510-2514
掲載論文のDOI (デジタルオブジェクト識別子) 10.1109/ISIT44484.2020.9173981	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 雲居玄道, 八木秀樹, 小林学, 後藤正幸, 平澤茂一	4. 巻 -
2. 論文標題 多値分類問題におけるECC法の最適性に関する一考察	5. 発行年 2021年
3. 雑誌名 情報処理学会論文誌数理モデル化と応用	6. 最初と最後の頁 -
掲載論文のDOI (デジタルオブジェクト識別子) なし	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 Umezawa Katsuyuki、Nakazawa Makoto、Kobayashi Manabu、Ishii Yutaka、Nakano Michiko、Hirasawa Shigeichi	4. 巻 -
2. 論文標題 Analysis of Logic Errors Utilizing a Large Amount of File History During Programming Learning	5. 発行年 2020年
3. 雑誌名 Proc. of IEEE Tale 2020	6. 最初と最後の頁 630-634
掲載論文のDOI (デジタルオブジェクト識別子) 10.1109/TALE48869.2020.9368348	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 S. Hirasawa, G. Kumoi, H. Yagi, M. Kobayashi, M. Goto, T. Sakai, and H. Inazumi	4. 巻 -
2. 論文標題 System Evaluation of Ternary Error-Correcting Output Codes for Multiclass Classification Problems	5. 発行年 2019年
3. 雑誌名 2019 IEEE International Conference on Systems, Man and Cybernetics	6. 最初と最後の頁 -
掲載論文のDOI (デジタルオブジェクト識別子) 10.1109/SMC.2019.8914295	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 S.Horii, T.Yoshida, M.Kobayashi, T.Matsushima	4. 巻 -
2. 論文標題 Distributed Stochastic Gradient Descent Using LDGM Codes	5. 発行年 2019年
3. 雑誌名 Proceedings of 2019 IEEE International Symposium on Information Theory (ISIT2019)	6. 最初と最後の頁 1417 - 1421
掲載論文のDOI (デジタルオブジェクト識別子) 10.1109/ISIT.2019.8849580	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 K. Umezawa, T. Ishida, M. Nakazawa, S.Hirasawa	4. 巻 -
2. 論文標題 Application of grouped flipped classroom to two-year actual class and its statistical evaluation	5. 発行年 2019年
3. 雑誌名 The 14th International Conference on Computer Science & Education	6. 最初と最後の頁 283-288
掲載論文のDOI (デジタルオブジェクト識別子) なし	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 山口純輝、風間阜希、鎌塚明、齋藤翔太、松嶋敏泰	4. 巻 103-A
2. 論文標題 拡張直交配列を用いた混合水準の実験計画法に関する一考察	5. 発行年 2020年
3. 雑誌名 電子情報通信学会論文誌(A)	6. 最初と最後の頁 17-24
掲載論文のDOI (デジタルオブジェクト識別子) なし	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 J. YOSHIKAWA, S. SAITO, T. MATSUSHIMA	4. 巻 E102-A
2. 論文標題 Variable-Length Intrinsic Randomness on Two Performance Criteria Based on Variational Distance	5. 発行年 2019年
3. 雑誌名 IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences	6. 最初と最後の頁 1642-1650
掲載論文のDOI (デジタルオブジェクト識別子) 10.1587/transfun.E102.A.1642	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 Mikawa Kenta, Kobayashi Manabu, Sasaki Tomoyuki, Manada Akiko	4. 巻 -
2. 論文標題 A Study on Latent Structural Models for Binary Relational Data with Attribute Information	5. 発行年 2023年
3. 雑誌名 Proceedings of 2023 International Symposium on Nonlinear Theory and Its Applications.	6. 最初と最後の頁 198-201
掲載論文のDOI (デジタルオブジェクト識別子) 10.34385/proc.76.A5L-411	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -

1. 著者名 Mikawa Kenta, Kobayashi Manabu, Sasaki Tomoyuki, Manada Akiko	4. 巻 15
2. 論文標題 A study on latent structural models for binary relational data with attribute information	5. 発行年 2024年
3. 雑誌名 Nonlinear Theory and Its Applications, IEICE	6. 最初と最後の頁 335 ~ 353
掲載論文のDOI (デジタルオブジェクト識別子) 10.1587/nolta.15.335	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -

1. 著者名 鈴木 佐俊、小林 学、後藤 正幸	4. 巻 74
2. 論文標題 世帯主幹消費電力量からのベイズ推定による個別家電の稼働状態の推定モデル	5. 発行年 2023年
3. 雑誌名 日本経営工学会論文誌	6. 最初と最後の頁 63 ~ 76
掲載論文のDOI (デジタルオブジェクト識別子) 10.11221/jima.74.63	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -

1. 著者名 NAKAHARA Yuta、MATSUSHIMA Toshiyasu	4. 巻 E107.A
2. 論文標題 Batch Updating of a Posterior Tree Distribution Over a Meta-Tree	5. 発行年 2024年
3. 雑誌名 IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences	6. 最初と最後の頁 523 ~ 525
掲載論文のDOI (デジタルオブジェクト識別子) 10.1587/transfun.2023TAL0003	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -

1. 著者名 Ishiwatari Taisuke、Saito Shota、Nakahara Yuta、Iikubo Yuji、Matsushima Toshiyasu	4. 巻 -
2. 論文標題 Bayes optimal estimation and its approximation algorithm for difference with and without treatment under IRSLC model	5. 発行年 2023年
3. 雑誌名 International Journal of Data Science and Analytics	6. 最初と最後の頁 -
掲載論文のDOI (デジタルオブジェクト識別子) 10.1007/s41060-023-00468-8	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

〔学会発表〕 計11件 (うち招待講演 0件 / うち国際学会 0件)

1. 発表者名 野口峻太, 須子統太, 小林学
2. 発表標題 不良回答検出を目的としたアンケート設計手法と検出率の理論評価について
3. 学会等名 情報処理学会第85回全国大会
4. 発表年 2023年

1. 発表者名 小林 学, 中原悠太, 佐藤絵里子, 高谷親信, 塚本好宏
2. 発表標題 銀行データを活用するための産学連携研究プラットフォームの構築
3. 学会等名 情報処理学会第84回全国大会
4. 発表年 2022年

1. 発表者名 小林 学, 松嶋敏泰, 平澤茂一
2. 発表標題 回帰木のベイズ的アプローチに基づくGAM/GA2M
3. 学会等名 日本経営工学会春季大会
4. 発表年 2021年

1. 発表者名 雲居玄道, 平澤茂一, 八木秀樹, 小林 学, 後藤正幸
2. 発表標題 構成的符号化を用いたECOCの一構成法
3. 学会等名 日本経営工学会春季大会
4. 発表年 2021年

1. 発表者名 良川太河, 小林 学, 後藤正幸
2. 発表標題 ベイズ最適化に基づく最適パラメータ探索法の研究動向と課題に関する一考察
3. 学会等名 情報処理学会第84回全国大会
4. 発表年 2022年

1. 発表者名 小林 学, 松嶋敏泰, 平澤茂一
2. 発表標題 回帰木のベイズ的アプローチに基づくGAM/GA2M
3. 学会等名 日本経営工学会春季大会
4. 発表年 2020年～2021年

1. 発表者名 小林 学, 雲居 玄道, 八木 秀樹, 平澤 茂一
2. 発表標題 ECOC法に基づく多値分類の平均的性能について
3. 学会等名 第46回情報理論とその応用シンポジウム
4. 発表年 2023年

1. 発表者名 須子統太, 小林学
2. 発表標題 不良回答の検出を目的としたアンケート設計に関する一考察
3. 学会等名 第46回情報理論とその応用シンポジウム
4. 発表年 2023年

1. 発表者名 石田崇, 雲居玄道, 小林学, 平澤茂一
2. 発表標題 生成 AI を用いた統計学の学習用練習問題自動生成の試み
3. 学会等名 日本教育工学会 2023年秋季全国大会
4. 発表年 2023年

1. 発表者名 石田崇, 雲居玄道, 小林学, 平澤茂一
2. 発表標題 生成 AI による統計学の学習用練習問題の品質の検証
3. 学会等名 日本教育工学会 2024年春季全国大会
4. 発表年 2024年

1. 発表者名 雲居玄道, 石田崇, 小林学, 平澤茂一
2. 発表標題 生成 AI による統計学の学習用練習問題の品質の検証
3. 学会等名 日本教育工学会 2024年春季全国大会
4. 発表年 2024年

〔図書〕 計3件

1. 著者名 松嶋 敏泰、早稲田大学データ科学教育チーム	4. 発行年 2023年
2. 出版社 サイエンス社	5. 総ページ数 192
3. 書名 データ科学入門II	

1. 著者名 松嶋敏泰, 中原悠太, 小林学, 堀井俊佑, 野村亮	4. 発行年 2022年
2. 出版社 サイエンス社	5. 総ページ数 192
3. 書名 データ科学入門I	

1. 著者名 松嶋敏泰, 堀井俊佑, 中原悠太, 小林学	4. 発行年 2024年
2. 出版社 サイエンス社	5. 総ページ数 176
3. 書名 データ科学入門III: モデルの候補が複数あるときの意思決定	

〔産業財産権〕

〔その他〕

早稲田大学研究者データベース
https://w-rdb.waseda.jp/html/100000011_ja.html

6. 研究組織

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
研究分担者	平澤 茂一 (Hirasawa Shigeichi) (30147946)	早稲田大学・理工学術院・名誉教授 (32689)	
研究分担者	松嶋 敏泰 (Matsushima Toshiyasu) (30219430)	早稲田大学・理工学術院・教授 (32689)	

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関
---------	---------