

令和 6 年 6 月 20 日現在

機関番号：10101

研究種目：基盤研究(B)（一般）

研究期間：2019～2023

課題番号：19H04128

研究課題名（和文）『稀』：気づきを与える認識・発見技術

研究課題名（英文）MARE: Recognition and Discovery Techniques for Giving Awareness

研究代表者

工藤 峰一（Kudo, Mineichi）

北海道大学・情報科学研究院・教授

研究者番号：60205101

交付決定額（研究期間全体）：（直接経費） 13,100,000円

研究成果の概要（和文）：希少疾患や独居高齢者の異変、さらには未発見の興味分野など、不頻出故に見過ごされがちなコトやモノにこそ重要な発見や意義があるとの立場で、不頻出事象の発見と分類を行った。不頻出事象の発生理由が、新規な事象であるほか、より細かなラベルが必要となったことに起因することが判明した。これらを踏まえて複数の分類手法を開発したものの、従来の性能を若干改善するに留まった。具体的応用としては、希少疾患を例に、通常事例との見分けに可視化技術が有効であることを示した。さらに、独居高齢者の異変検出のため、仮想スマートホームにおける行動シミュレータ及び異変検知手法を開発した。

研究成果の学術的意義や社会的意義

不頻出事象はこの世界にあまねく存在するものの、その希少性により扱いが難しい。本研究はそれらの分類についての方法論を深化させた。特に、専門家でも難しい希少疾患の発見の一助になる方法を提示できたこと、スマートホームを利用して独居高齢者の異変を検知する試みの基本的有効性を示せたことの意義は大きい。

研究成果の概要（英文）：On the standing point that rare events, such as rare diseases, anomaly of elderly living alone and interesting but unknown area, are more important than usual events, we have developed techniques for discovering rare events and classifying them. The reasons of occurrence of rare events were revealed to be "new" and devised from the necessity of finer labels. Accordingly, on the basis of these reasons, we developed some classifiers, but only a small amount of improvement was achieved by them.

As an application, on discovery of rare diseases, we have shown that visualization is useful for distinguishing rare events from ordinary events. For detecting anomalies of elderly living alone, we have developed a behavior simulator in a virtual smart home and detection algorithms of anomaly detection.

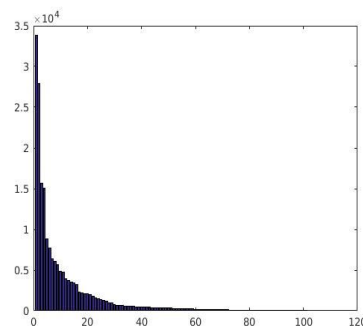
研究分野：パターン認識

キーワード：不頻出事象 マルチラベル分類 希少疾患の発見 高齢者の異変検知

## 様式 C - 19、F - 19 - 1 (共通)

### 1. 研究開始当初の背景

希少疾患や故障など、世の中には滅多に起きない事象や例外的なできごとが多く存在する。数や種類においてはむしろよく起きる一般的な事象よりも多いこともある。不頻出であることは、単に出現が稀であるという以外に、特定の(専門的)であることや、新規なものであるということに起因することが多く、そのようなコトやモノを知ることは新しい気づきや重要な発見につながる。しかし、一般的な不頻出事象の認識や推論に関しては、事例の少なさによる学習の困難さに加え、成功例数に基づく一般的な評価には貢献しないことからこれまで殆ど研究されていない。本研究では、希少疾患の可能性示唆など、不頻出なコトやモノの認識や発見を行う一般的枠組みをパターン認識およびデータマイニング技術により確立する。



ラベル ID  
図 1 ラベルの頻度分布  
(Mediamill dataset)

### 2. 研究の目的

これまで困難であった不頻出事象の分析と予測を理論面で検討するとともに各種方法論を実践することを目的とする。理論面では、主にパターン認識における(エクストリーム)マルチラベル分類の一つとして検討を行う。不頻出ラベルはロングテール分布の裾にあり、その希少性(インバランス)から通常と別な方法論の構築が必要である。本研究では、不頻出ラベルの発生原因の分析から始め、分類手法の確立を目指す。応用面では、希少疾患の分類および独居高齢者の異変検知を行う。希少疾患を経験の浅い医師でも見分けられるような補助手段を開発する。独居高齢者においては典型的な異変をスマートホームにおいて自動で検出する方式を開発する。

### 3. 研究の方法

理論面での成果と応用面での方法論を順に述べる。

#### 【理論面】

- ・(分析)(エクストリーム)マルチラベル分類のデータベースを利用してロングテール分布を調べ、ヘッド部分とテール部分の相関を調べてテール部分にある希少ラベルの予測可能性を調べる。また、可視化技法を利用した分析も行う。
- ・(分類) サンプルバランスの均一化を重視する手法とラベルに依存して有効な特徴を選出する手法の2手法を検討する。

#### 【応用面】

- ・(希少疾患の分類) 希少疾患データベースを利用して通常疾患との相違点を調べる。次に、この相違点に基づいて分類器を構成する。超矩形を用いた方式と決定木(森)を用いた方式を基に希少疾患の検出率を重視した方法を構築する。分析として、可視化により、特定の疾患がどの通常疾患に間違われやすいか、また、それは何故かを調べる。
- ・(独居高齢者の異変検出) 高齢者の異変は数秒間の変化である転倒から数か月に渡る変化である準寝たきりなど多種多様である。そこで環境センサを配置したスマートホームにおいて典型的な異変を検出する方法論を確立する。現実の異変データを獲得することはほぼ不可能なので仮想空間でこれを行う。

### 4. 研究成果

理論面での成果と応用面での成果を順に述べる。

#### 【理論面】

- ・(発生原因と難しさの源泉) データセットに基づいて希少ラベル(テールラベル)の発生原因を調べた。その結果、主な発生理由は、1) 新規ラベルとして新たに認知されたため、2) これまでよりも精密な分類が必要となり創出されたため、の二つであることが分かった。また、分類の難しさは、事例の少なさ(インバランス)ばかりでなく、2)の発生理由からもわかるように、多くの類似データに埋もれやすいため(オーバーラップ)であることが判明した。方法論として、希少ラベルを強調した可視化による希少例の分析の有効性を示した。
- ・(ロングテール分類への挑戦) 超区間と決定木を用いた新たな分類手法を複数提案した。特徴選択の有効性と階層化による実用的な分類の効果を確認したものの、希少ラベルの再現率を上げる本質的な改善にはつながらなかった。
- ・(可視化による分析) ラプラシアン固有マップをマルチラベルに対応できるように改良した。加えて、教師付きに対応するように改良した。これらの非線形「埋め込み」手法は未知のサン

ルを同じ方式で可視化できないという問題がある。そこで、カーネルの線形和を用いてこの方式を模擬する方式を開発しその有効性をパラメータ選択で示した。

### 【応用面】

・(希少疾患の分類) 極値理論に基づく分類方式を開発して分類性能の改善を図った。また、先に開発した可視化技法を用いて、特定の希少疾患がどの通常疾患に間違われている可能性が高いかを示すことで審査医を補助する方式を提案した。

### ・(独居高齢者の異変検出)

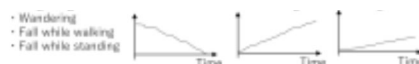
独居高齢者の異変検出については、7つ程度の典型的な異変をスマートホーム環境で検出することを考えた。現実の異変データはほぼ入手困難であるため、仮想空間でのシミュレーションにより異変検出を行うこととした。

そのため、異変を含む独居高齢者の日常生活を模擬するシミュレータを開発し数年間の行動シミュレーションを数十分で行えるようにした。転倒についてはこれまで以上

に多様な生活環境でも同程度の発見を行える手法を開発した。現在、現実とシミュレーションとの乖離を埋めつつ他の異変を併せて検出する方法を開発している。



図2 行動シミュレーション。上から、部屋を移動する高齢者、その移動に対応したスケジュール、6種類の異変と認知症の進展に応じた異変頻度の増加グラフ



5. 主な発表論文等

〔雑誌論文〕 計12件（うち査読付論文 12件 / うち国際共著 3件 / うちオープンアクセス 2件）

1. 著者名 Tai Mariko, Kudo Mineichi, Tanaka Akira, Imai Hideyuki, Kimura Keigo	4. 巻 123
2. 論文標題 Kernelized Supervised Laplacian Eigenmap for Visualization and Classification of Multi-Label Data	5. 発行年 2022年
3. 雑誌名 Pattern Recognition	6. 最初と最後の頁 108399 ~ 108399
掲載論文のDOI (デジタルオブジェクト識別子) 10.1016/j.patcog.2021.108399	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 Mineichi Kudo, Keigo Kimura, Shumpei Morishita and Lu Sun	4. 巻 13813
2. 論文標題 Efficient Leave-One-Out Evaluation of Kernelized Implicit Mappings	5. 発行年 2023年
3. 雑誌名 Lecture Notes in Computer Science	6. 最初と最後の頁 223-232
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 該当する

1. 著者名 Kimura Keigo, Bao Jiaqi, Kudo Mineichi and Sun Lu	4. 巻 13813
2. 論文標題 Retargeted Regression Methods for Multi-label Learning	5. 発行年 2023年
3. 雑誌名 Lecture Notes in Computer Science	6. 最初と最後の頁 203-212
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 該当する

1. 著者名 Shumpei Morishita, Mineichi Kudo, Keigo Kimura, and Lu Sun	4. 巻 13813
2. 論文標題 Realization of Autoencoders by Kernel Methods	5. 発行年 2023年
3. 雑誌名 Lecture Notes in Computer Science	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 該当する

1. 著者名 Tanaka Kai, Kudo Mineichi, Kimura Keigo	4. 巻 2022
2. 論文標題 Sensor Data Simulation with Wandering Behavior for the Elderly Living Alone	5. 発行年 2022年
3. 雑誌名 Proc. of the 26TH International Conference on Pattern Recognition	6. 最初と最後の頁 885-891
掲載論文のDOI (デジタルオブジェクト識別子) 10.1109/ICPR56361.2022.9956332	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 LU Kejing, KUDO Mineichi	4. 巻 E104.D
2. 論文標題 AdaLSH: Adaptive LSH for Solving $\epsilon$ -Approximate Maximum Inner Product Search Problem	5. 発行年 2021年
3. 雑誌名 IEICE Transactions on Information and Systems	6. 最初と最後の頁 138 ~ 145
掲載論文のDOI (デジタルオブジェクト識別子) 10.1587/transinf.2020EDP7132	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 Lu Kejing, Kudo Mineichi, Xiao Chuan, Ishikawa Yoshiharu	4. 巻 15
2. 論文標題 HVS	5. 発行年 2021年
3. 雑誌名 Proceedings of the VLDB Endowment	6. 最初と最後の頁 246 ~ 258
掲載論文のDOI (デジタルオブジェクト識別子) 10.14778/3489496.3489506	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 Aoki Shuhei, Kudo Mineichi	4. 巻 13055
2. 論文標題 Balancing of Samples in Class Hierarchy	5. 発行年 2021年
3. 雑誌名 Lecture Notes in Computer Science	6. 最初と最後の頁 219 ~ 228
掲載論文のDOI (デジタルオブジェクト識別子) 10.1007/978-3-030-89691-1_22	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 Horio Tomoya, Kudo Mineichi	4. 巻 13055
2. 論文標題 Feature Selection with Class Hierarchy for Imbalance Problems	5. 発行年 2021年
3. 雑誌名 Lecture Notes in Computer Science	6. 最初と最後の頁 229 ~ 238
掲載論文のDOI (デジタルオブジェクト識別子) 10.1007/978-3-030-89691-1_23	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 Kaneko Yasuyuki, Kudo Mineichi	4. 巻 13055
2. 論文標題 SVM Based EVM for Open Space Problems	5. 発行年 2021年
3. 雑誌名 Lecture Notes in Computer Science	6. 最初と最後の頁 239 ~ 248
掲載論文のDOI (デジタルオブジェクト識別子) 10.1007/978-3-030-89691-1_24	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 Lu Kejing, Kudo Mineichi	4. 巻 12682
2. 論文標題 MLSH: Mixed Hash Function Family for Approximate Nearest Neighbor Search in Multiple Fractional Metrics	5. 発行年 2021年
3. 雑誌名 Lecture Notes in Computer Science	6. 最初と最後の頁 569 ~ 584
掲載論文のDOI (デジタルオブジェクト識別子) 10.1007/978-3-030-73197-7_38	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 Lu Kejing, Wang Hongya, Wang Wei, Kudo Mineichi	4. 巻 13
2. 論文標題 VHP: Approximate Nearest Neighbor search via Virtual Hypersphere partitioning	5. 発行年 2020年
3. 雑誌名 Proceedings of the VLDB Endowment	6. 最初と最後の頁 1443 ~ 1455
掲載論文のDOI (デジタルオブジェクト識別子) 10.14778/3397230.3397240	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18件(うち招待講演 0件/うち国際学会 14件)

1. 発表者名 Mineichi Kudo, Keigo Kimura, Shumpei Morishita and Lu Sun
2. 発表標題 Efficient Leave-One-Out Evaluation of Kernelized Implicit Mappings
3. 学会等名 S+SSPR2022 (国際学会)
4. 発表年 2022年

1. 発表者名 Kimura Keigo, Bao Jiaqi, Kudo Mineichi and Sun Lu
2. 発表標題 Retargeted Regression Methods for Multi-label Learning
3. 学会等名 S+SSPR2022 (国際学会)
4. 発表年 2022年

1. 発表者名 Shumpei Morishita, Mineichi Kudo, Keigo Kimura, and Lu Sun
2. 発表標題 Realization of Autoencoders by Kernel Methods
3. 学会等名 S+SSPR2022 (国際学会)
4. 発表年 2022年

1. 発表者名 Tanaka Kai, Kudo Mineichi, Kimura Keigo
2. 発表標題 Sensor Data Simulation with Wandering Behavior for the Elderly Living Alone
3. 学会等名 ICPR2022 (国際学会)
4. 発表年 2022年

1. 発表者名 Lu Kejing, Kudo Mineichi
2. 発表標題 MLSH: Mixed Hash Function Family for Approximate Nearest Neighbor Search in Multiple Fractional Metrics
3. 学会等名 26th International Conference (DASFAA 2021) (国際学会)
4. 発表年 2021年

1. 発表者名 Aoki Shuhei, Kudo Mineichi
2. 発表標題 Balancing of Samples in Class Hierarchy
3. 学会等名 7th International Workshop on Artificial Intelligence and Pattern Recognition (IWAIPR2021) (国際学会)
4. 発表年 2021年

1. 発表者名 Kaneko Yasuyuki, Kudo Mineichi
2. 発表標題 SVM Based EVM for Open Space Problems
3. 学会等名 7th International Workshop on Artificial Intelligence and Pattern Recognition (IWAIPR2021) (国際学会)
4. 発表年 2021年

1. 発表者名 Horio Tomoya, Kudo Mineichi
2. 発表標題 Feature Selection with Class Hierarchy for Imbalance Problems
3. 学会等名 7th International Workshop on Artificial Intelligence and Pattern Recognition (IWAIPR2021) (国際学会)
4. 発表年 2021年



1. 発表者名 Lu Kejing, Kudo Mineichi
2. 発表標題 R2LSH: A Nearest Neighbor Search Scheme Based on Two-dimensional Projected Spaces
3. 学会等名 ICDE 2020 ( 国際学会 )
4. 発表年 2020年

1. 発表者名 Mitsuki Maekawa, Atsuyoshi Nakamura, Mineichi Kudo
2. 発表標題 Data-Dependent Conversion to a Compact Integer-Weighted Representation of a Weighted Voting Classifier
3. 学会等名 Proceedings of The 12th Asian Conference on Machine Learning ( 国際学会 )
4. 発表年 2020年

1. 発表者名 Mariko Tai and Mineichi Kudo
2. 発表標題 A Supervised Laplacian Eigenmaps Algorithm for Visualization of Multi-label Data: SLE-ML
3. 学会等名 24th Iberoamerican Congress on Pattern Recognition ( 国際学会 )
4. 発表年 2019年

1. 発表者名 M Maekawa, A Nakamura, M Kudo
2. 発表標題 Conversion from a Real-Weighted Majority Voting Classifier to a Small-Non-Negative-Integer-Weighted Thresholded Voting Classifier
3. 学会等名 ACML 2019 Workshop on Statistics & Machine Learning Researchers in Japan ( 国際学会 )
4. 発表年 2019年

1. 発表者名 K Todo, A Nakamura, M Kudo
2. 発表標題 A Fast Approximate Algorithm for k-Median Problem on a Graph.
3. 学会等名 15th International Workshop on Mining and Learning with Graphs (国際学会)
4. 発表年 2019年

1. 発表者名 R Shirakawa, A Nakamura, M Kudo
2. 発表標題 Learning a Nonlinear Model of Subgraph Features Using Monte Carlo Tree Search
3. 学会等名 ACML 2019 Workshop on Statistics & Machine Learning Researchers in Japan (国際学会)
4. 発表年 2019年

1. 発表者名 横山祐也, 工藤峰一
2. 発表標題 ルールアンサンブル法を用いたマイノリティクラスの識別,
3. 学会等名 電子情報通信学会技術研究報告, PRMU
4. 発表年 2019年

1. 発表者名 田井茉莉子, 工藤峰一
2. 発表標題 マルチラベル判別ラプラシアン固有マップのカーネル化
3. 学会等名 電子情報通信学会技術研究報告, PRMU-59
4. 発表年 2019年

1. 発表者名 小城佑樹, 工藤峰一
2. 発表標題 微小なうなずきの検出精度向上
3. 学会等名 電子情報通信学会技術研究報告, PRMU-59
4. 発表年 2019年

1. 発表者名 伊藤直輝, 中村篤祥, 工藤峰一
2. 発表標題 ブラックボックス関数のノイズ入り関数値からの閾値以上の値の存在チェック
3. 学会等名 第22回情報論的学習理論ワークショップ
4. 発表年 2019年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

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	今井 英幸  (Imai Hideyuki)  (10213216)	北海道大学・情報科学研究院・教授   (10101)	
研究 分担者	中村 篤祥  (Nakamura Atsuyoshi)  (50344487)	北海道大学・情報科学研究院・教授   (10101)	

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
---------	---------