リサーチペーパー第 63 号　　Research Paper No.63

収支項目分類符号に対する効率的な自動格付に関する研究

Autocoding for labels of Income and Expenditure classification

床　裕佳子
独立行政法人統計センター統計編成部消費統計編成課
TOKO Yukako
National Statistics Center
Statistical Data Processing Department
Consumer Statistics Data Processing Division

佐藤-イリチュ　美佳
統計研究研修所客員教授
筑波大学教授
SATO-ILIC Mika
SRTI Guest Professor
Professor, University of Tsukuba

令和 7 年 6 月
June 2025

総務省統計研究研修所
Statistical Research and Training Institute (SRTI)
Ministry of Internal Affairs and Communications

受理日：令和 7 年 5 月 22 日

本ペーパーは、総務省統計研究研修所の客員教授及び独立行政法人統計センター職員である執筆者が、その責任において行った統計研究の成果を取りまとめたものであり、その内容については、統計研究研修所または統計センターの見解を表したものではない。本ペーパーの内容については、執筆者に問い合わせ願いたい。

本研究では、統計法(平成 19 年法律第 53 号)第 32 条及び 33 条の規定に基づき、家計調査、全国家計構造調査及び全国単身世帯収支実態調査に係る調査票情報を使用した。

収支項目分類符号に対する自動格付

床　裕佳子
佐藤-イリチュ　美佳

概要

独立行政法人統計センターでは、「家計調査」における各調査月のオンライン回答された家計簿の各収入項目・支出項目に対するデータについて、収支品目分類符号の自動付与による格付業務支援を目的として、機械学習型格付手法を用いたシステムを開発し、2022 年 1 月から運用が開始されている。本稿では、格付支援にあたって開発した方法を紹介する。

キーワード：格付、ファジィクラスタリング、サポートベクトルマシン、T-norm

Autocoding for labels of Income and Expenditure classification

TOKO Yukako
SATO-ILIC Mika

Abstract

The National Statistics Center has developed a system using a machine learning based autocoding method to support coding operations. It works by automatically assigning income and expenditure classification codes to data on each income and expenditure item in the household ledger. This ledger was answered online for each survey month in the "Family Income and Expenditure Survey." The system began operation in January 2022. This paper introduces the methods developed for autocoding support.

Keywords：Autocoding, Fuzzy clustering, Support vector machine, T-norms

1．はじめに

独立行政法人統計センターでは、「家計調査」における各調査月のオンライン回答された家計簿の各収入項目・支出項目に対するデータについて、収支品目分類符号の自動付与による格付業務支援を目的として、ルールベース型格付支援システムと機械学習型格付支援システムを組み合わせたハイブリッド型格付支援システムを開発し、2022年 1 月から運用が開始されている。図 1 は、ハイブリッド型格付支援システムの格付対象であるオンラインデータにおける 2019 年 1 月～2024 年 12 月の各月のデータ件数の推移である。赤線はレシート以外のデータの格付対象のデータ件数、青線はレシートデータの格付対象のデータ件数である。

図 1 に示すとおり、ハイブリッド型格付支援システムの格付対象となるオンラインデータのデータ件数は、毎月増加傾向にあり、特にレシートデータについては顕著に格付対象となるデータ件数が増加していることがわかる。図 2 は、ハイブリッド型格付支援システムが導入された 2022 年 1 月～2024 年 12 月におけるハイブリッド型格付支援システムによる格付率の推移である。青線はレシートデータの格付率、赤線はレシート以外のデータの格付率である。この図から、レシートデータの格付率は徐々に向上しているが、レシート以外のデータの格付率は 2023 年の上半期頃から徐々に低下傾向にあることがわかる。一方、レシート以外のデータに対する格付率とレシートデータに対する格付率を比較した場合においては、レシートデータの格付率の方が低い傾向にある。これはレシートデータが、人間が入力する文字情報とは異なる特徴を持つことに起因する。このような状況下において、現行の格付支援システムによる格付手法では格付困難となるようなデータ構造をもつデータが今後ますます増加する可能性がある。そのため、いくつかの機械学習型格付手法について提案を行ってきた。本稿では、これらの方法の結果について、紹介する。

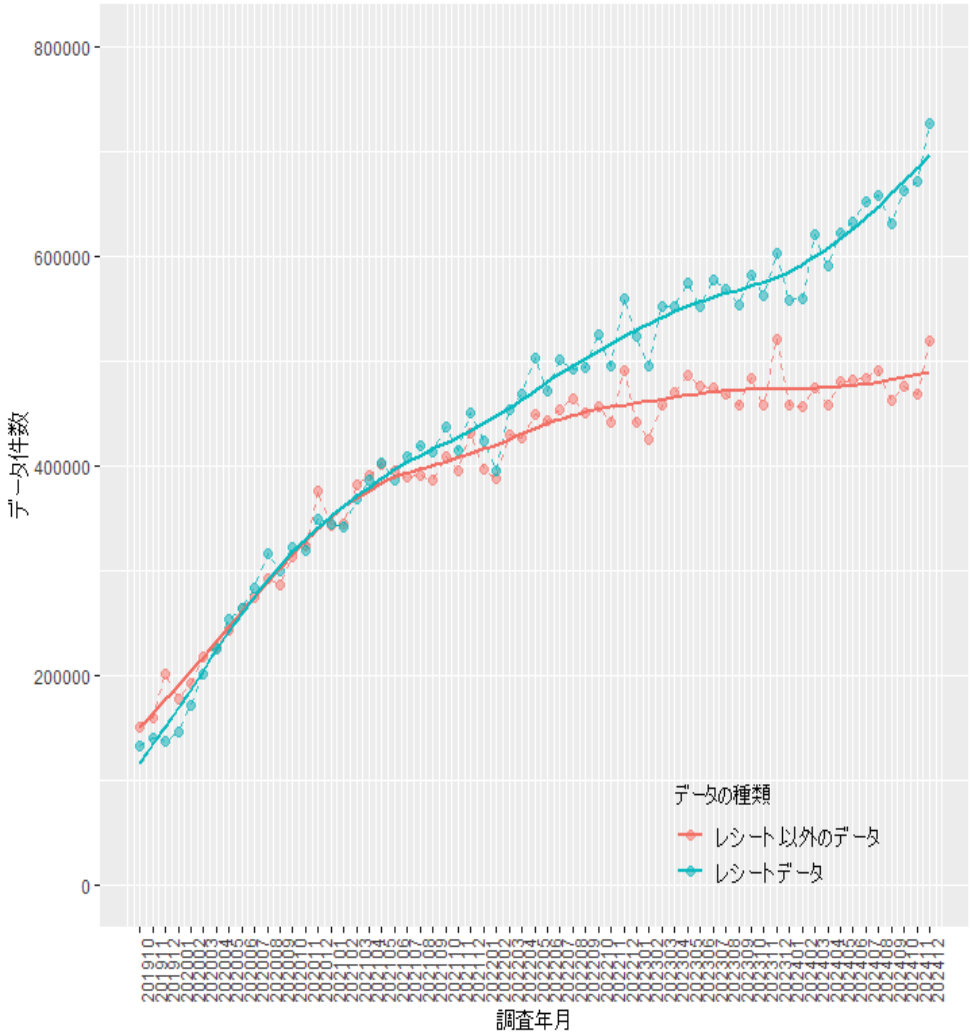


図 1. ハイブリッド型格付支援システムにおける格付対象データ件数の推移

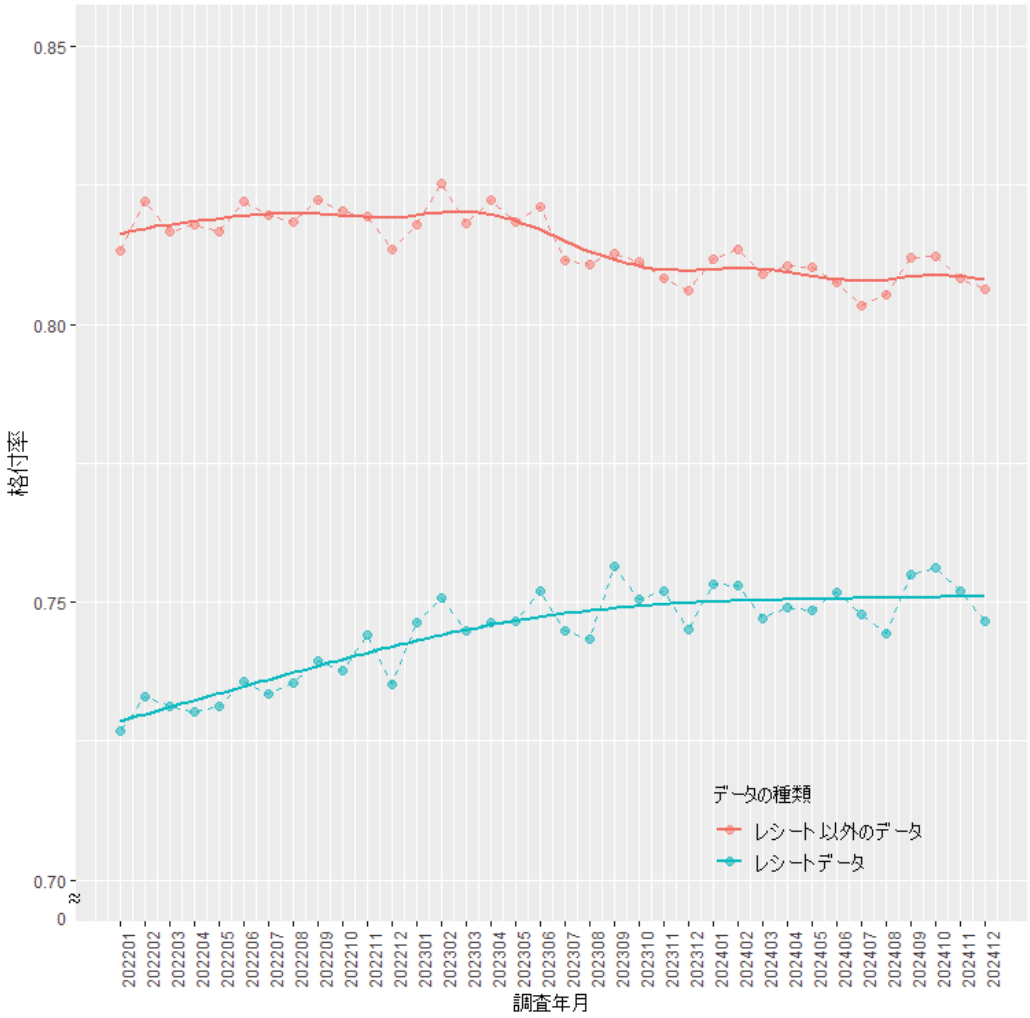


図 2. ハイブリッド型格付支援システムにおける格付率の推移

2．収支項目分類格付のためのハイブリッド型格付支援システムの概要

家計調査における家計簿の分類符号格付においては、オンラインによる回答が導入された 2019 年 1 月から、オンラインで回答された家計簿の格付業務において、格付率の向上を目指し、機械学習型格付手法として信頼度に基づく分類モデル[23]が開発され、その後、さらなる精度向上を目指し、T-norm を導入して一般化を図った方法を提案した[26]。 この機械学習型格付手法と既存のルールベース型格付手法を組み合せたハイブリッド型格付支援システムを開発し、2022 年 1 月から適用を開始した。ハイブリッド型格付支援システムでは、格付対象のデータに対して、まずルールベース型手法による格付を行い、その後、ルールベース型手法で格付できなかった未格付データに対して、機械学習型手法による格付を行う。

3．数値例

家計調査のオンラインデータを用いてハイブリッド型格付支援システムの性能評価を行った結果を示す。本検証で用いたデータには、日本語で入力された各収入項目またはレシートデータを含む支出品名及びそれらに対応する収支項目分類符号が含まれており、収支項目分類には約 520 の異なる符号が存在する。

図 3 は、2022 年 1 月～2024 年 12 月までのルールベース型格付手法の格付率の推移を示しており、図 4 は、同じ期間の機械学習型格付手法の格付率の推移を示している。図 3 及び図 4 の青線はレシートデータの格付率、赤線はレシート以外のデータの格付率である。図 3 から、ルールベース型の格付手法では、レシートデータ、レシート以外のデータ共に 2023 年の上半期頃から格付率が低下傾向にあることがわかる。一方で、図 4 から、機械学習型の格付手法では、レシートデータ、レシート以外のデータ共に格付率が向上しており、特にレシートデータにおいて大きく格付率が向上していることがわかる。図 2、図 3、図 4 から、レシートデータにおいては、機械学習型格付手法がルールベース型格付手法の格付率の低下分をカバーして格付していると言える。一方で、レシート以外のデータについては、機械学習型格付手法において格付率が向上しているが、ルールベースの格付率の低下分をカバーしきれておらず、全体として格付率が低下してきている。

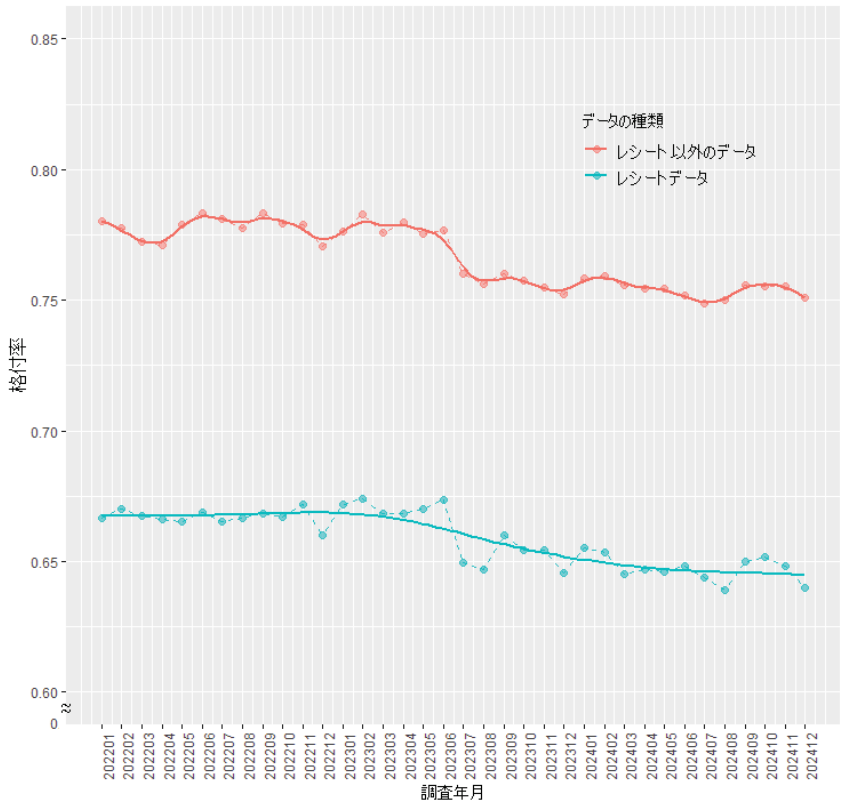


図 3. ルールベース型格付支援システムにおける格付率の推移

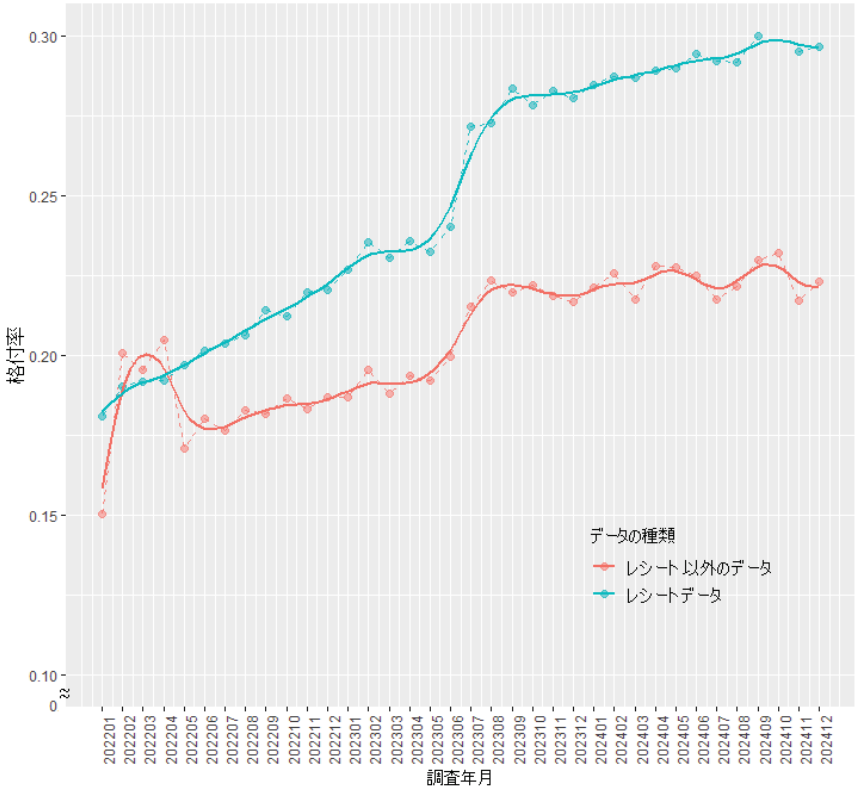


図 4. 機械学習型格付支援システムにおける格付率の推移

また、格付率のほか、正解率、マクロ適合率、マクロ再現率、マクロ f1 スコアの評価指標を用いて格付精度の検証を行った。

図 5 に、機械学習手法の信頼度にファジィ分類構造を示す尺度として分割係数、*T*-ノルムとして代数積を採用した場合のハイブリッド型格付支援システムにおける各評価指標の値及び格付率を示す。図 5 において、赤線で示した正解率は、評価データの全ての期間において、レシートデータ、レシート以外のデータ共に値が 0.99 を超えて安定して推移しているが、青線で示したマクロ適合率、紫線で示したマクロ再現率、緑線で示したマクロ f1 スコアについては、特にデータの種類（レシート以外のデータ/レシートデータ）によって値がばらついており安定しているとは言えない状況である。加えて、格付率が低下している現状から、格付対象のデータにおいてレシートデータをはじめ格付困難な複雑なデータが増加してきていると考えられ、このようなデータは今後ますます増加すると考えられる。また、他の信頼度を用いた場合においても概ね同様の傾向がみられた。図 6 ～図 13 に、2022 年 1 月～2023 年 6 月までの家計調査のオンラインデータを用いて、信頼度ごとのハイブリッド型格付支援システムの格付精度の比較を行った結果を示す。

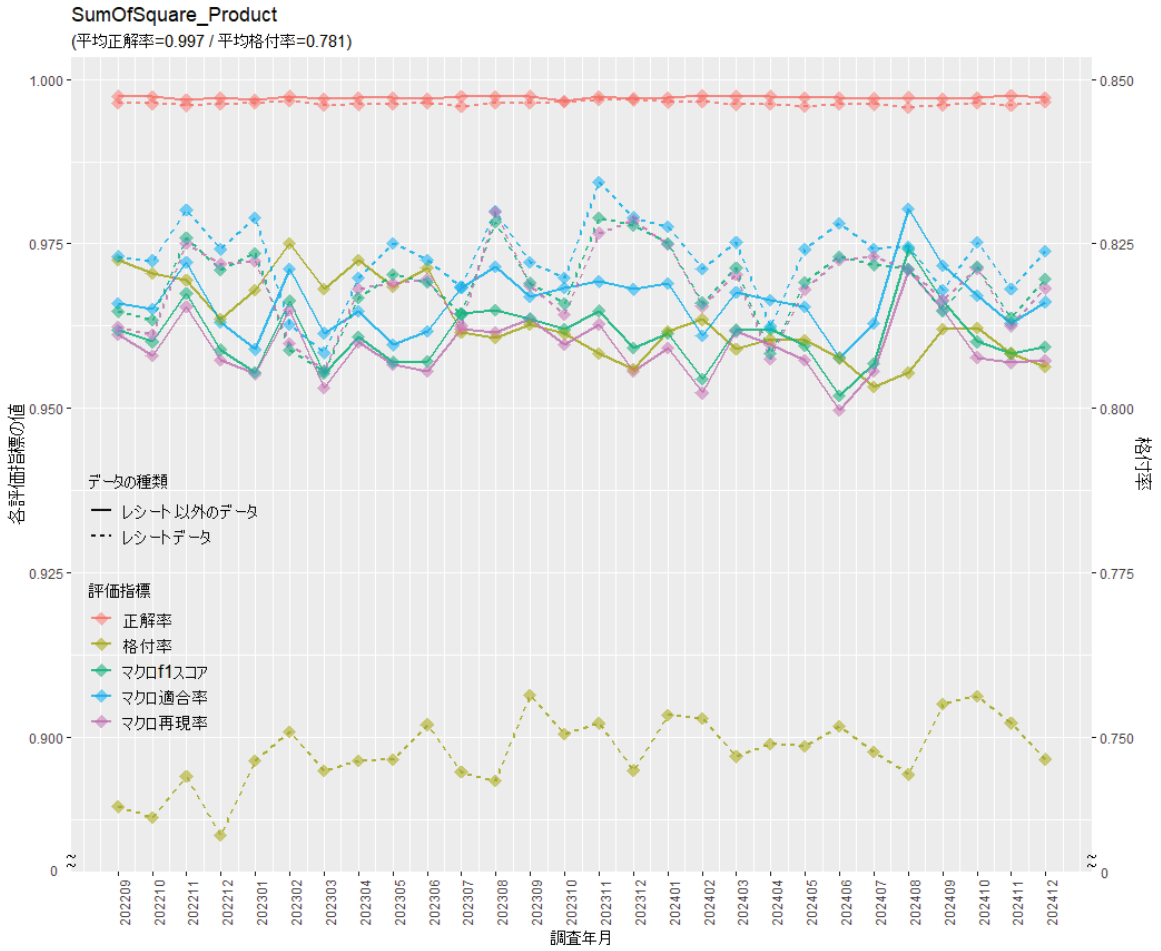


図 5. ハイブリッド型格付支援システムにおける各評価指標の値及び格付率

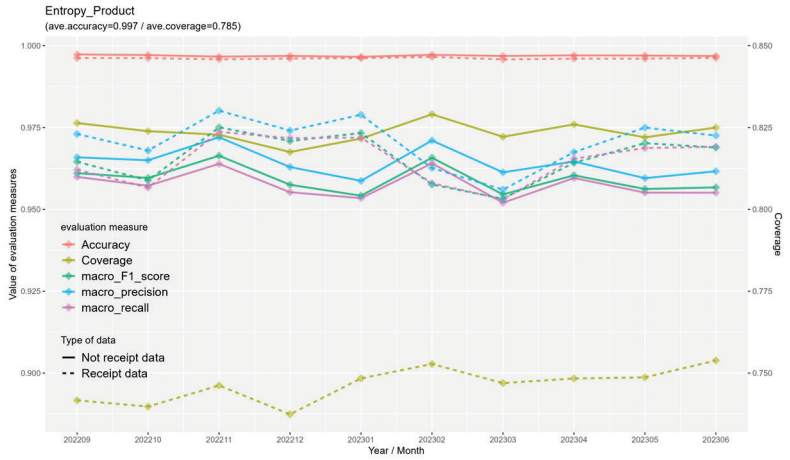


図 6. 分割エントロピーに基づく代数積を用いた信頼度の格付結果

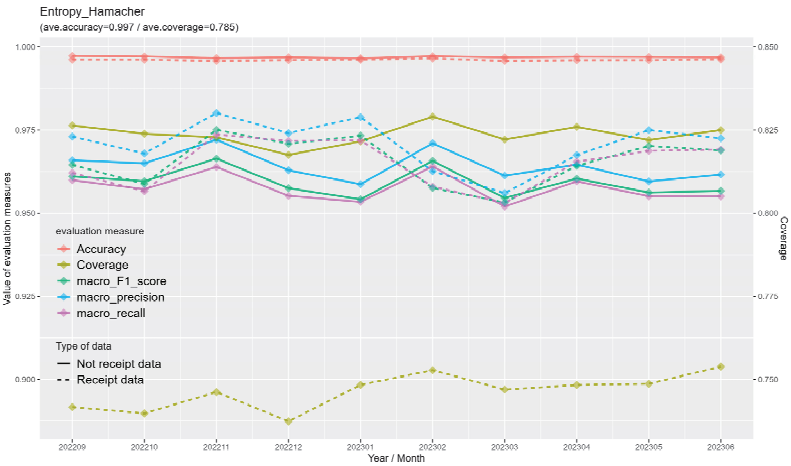


図 7. 分割エントロピーに基づくハマーカー積を用いた信頼度の格付結果

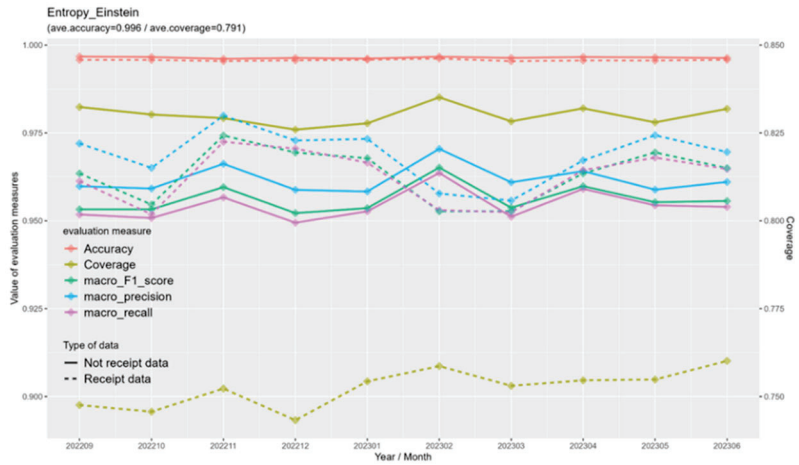


図 8. 分割エントロピーに基づくアインシュタイン積を用いた信頼度の格付結果

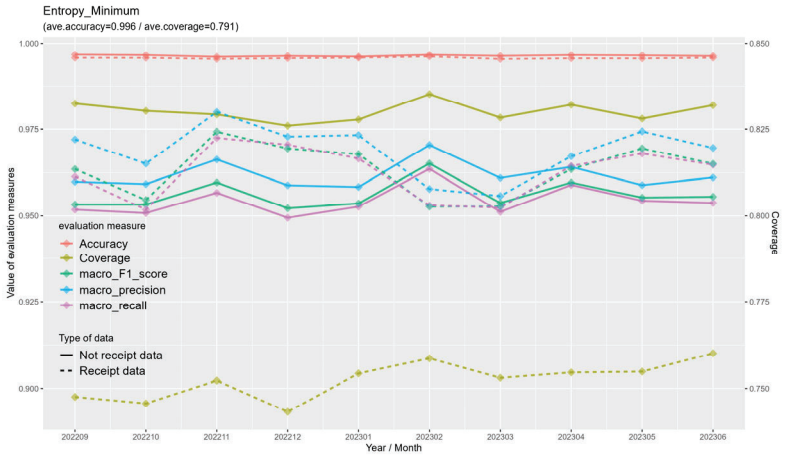


図 9. 分割エントロピーに基づく最小値を用いた信頼度の格付結果

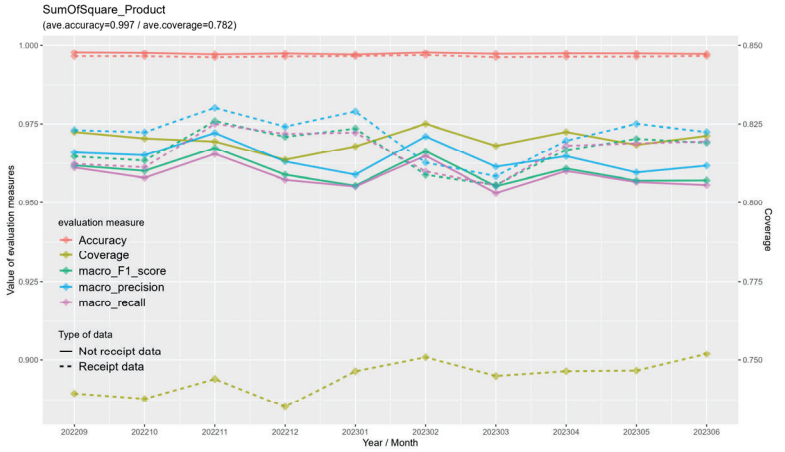


図 10. 分割係数に基づく代数積を用いた信頼度の格付結果

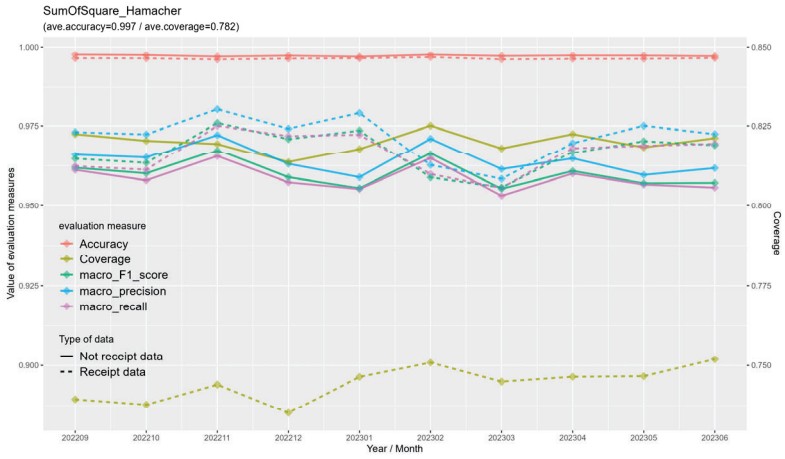


図 11. 分割係数に基づくハマーカー積を用いた信頼度の格付結果

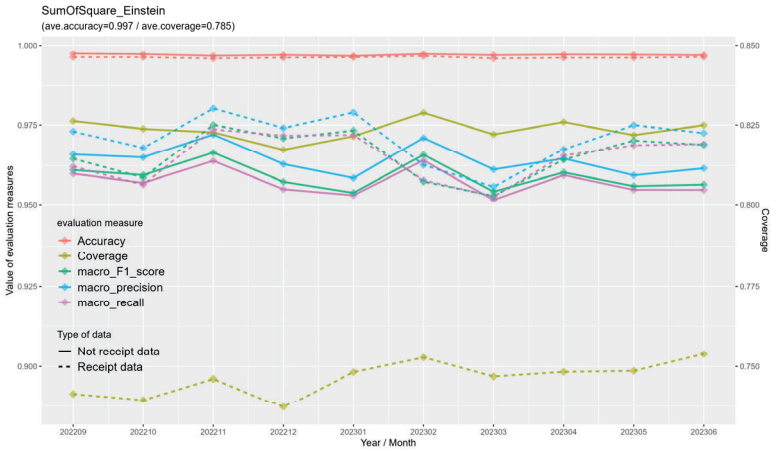


図 12. 分割係数に基づくアインシュタイン積を用いた信頼度の格付結果

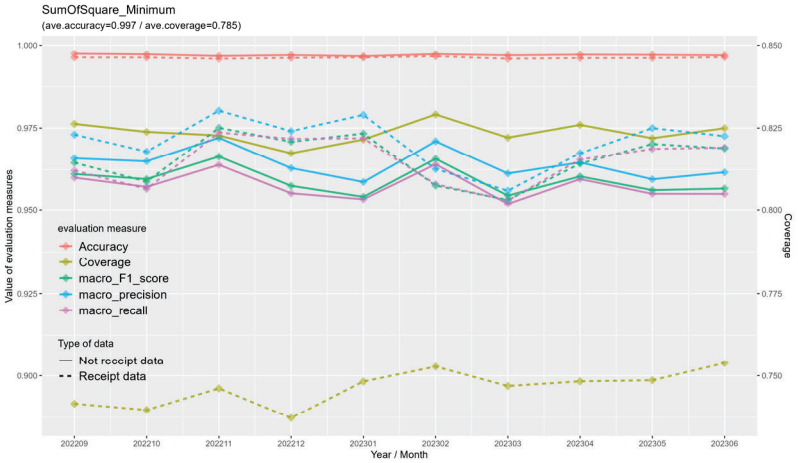
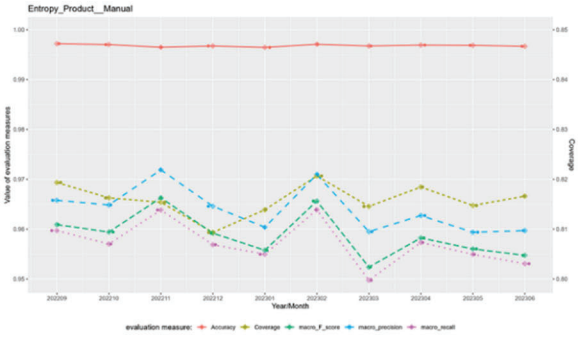
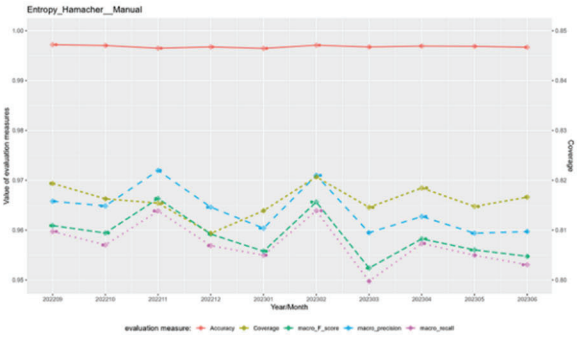


図 13. 分割係数に基づく最小値を用いた信頼度の格付結果

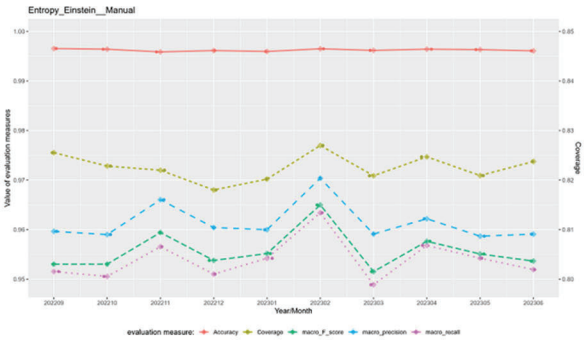
さらに、図 14～図 16 に示したとおり、より詳細な入力データの種類ごとの格付結果の傾向を調べるために、オンライン家計簿に手入力されたデータ、文字認識機能を用いて入力されたレシートデータ、オペレータにより入力されたレシートおよび明細書のデータについて、それぞれ信頼度ごとの格付精度を検証した。



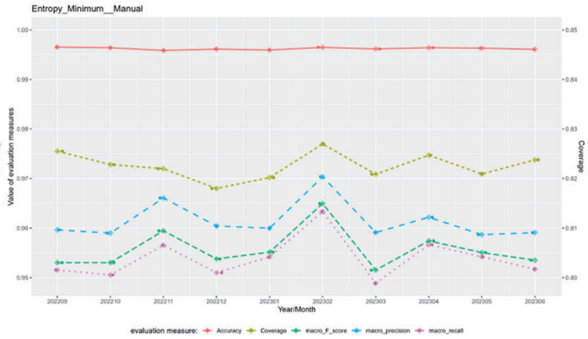
(a)信頼度：分割エントロピーに基づく代数積



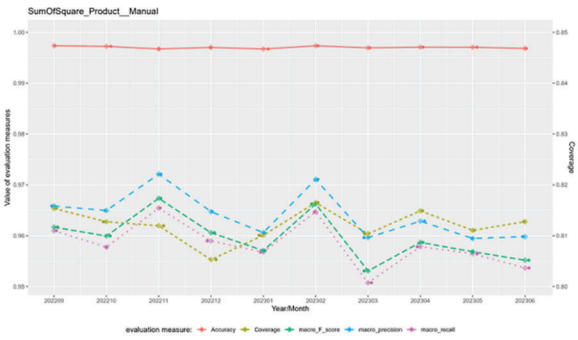
(b)信頼度：分割エントロピーに基づくハマーカー積



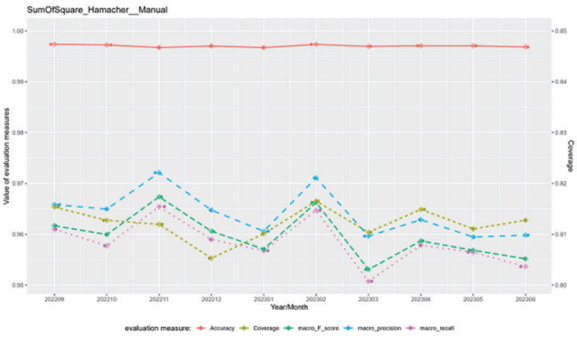
(c)信頼度：分割エントロピーに基づくアインシュタイン積



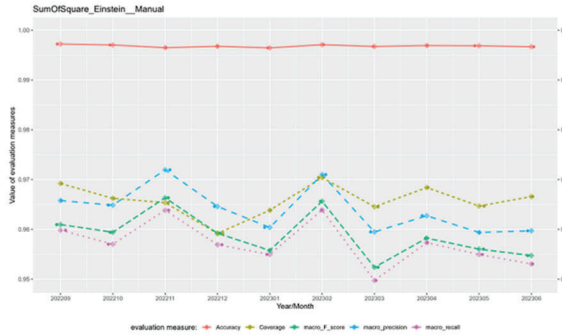
(d)信頼度：分割エントロピーに基づく最小値



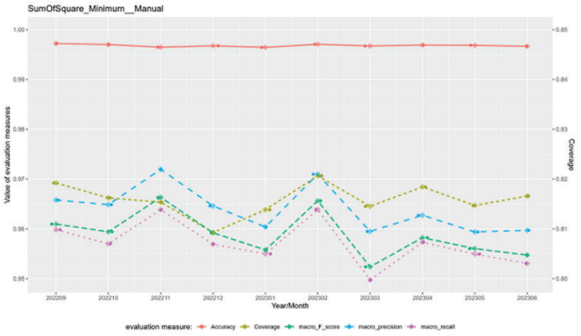
(e)信頼度：分割係数に基づく代数積



(f)信頼度：分割係数に基づくハマーカー積

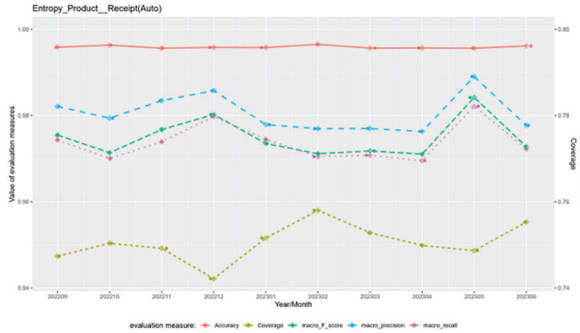


(g)信頼度：分割係数に基づくアインシュタイン積

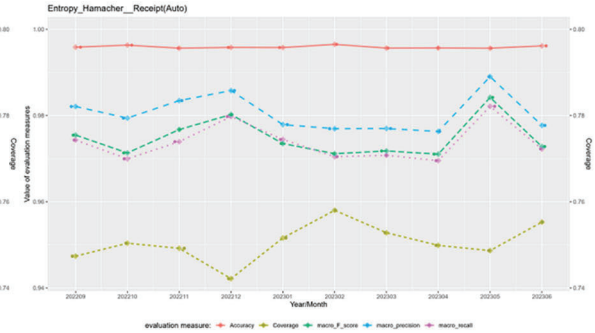


(h)信頼度：分割係数に基づく最小値

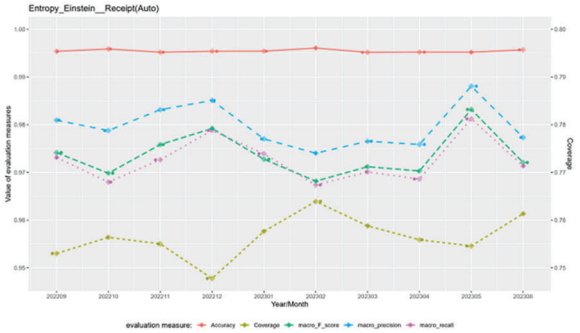
図 14. 手入力データにおける信頼度ごとの格付結果



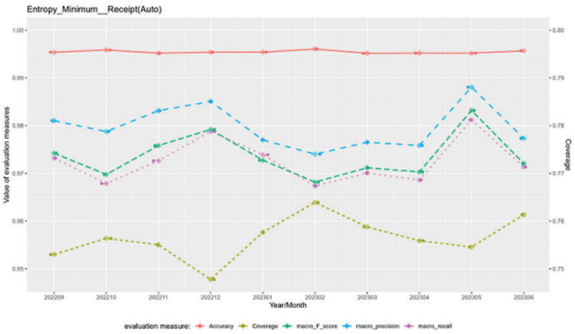
(a)信頼度：分割エントロピーに基づく代数積



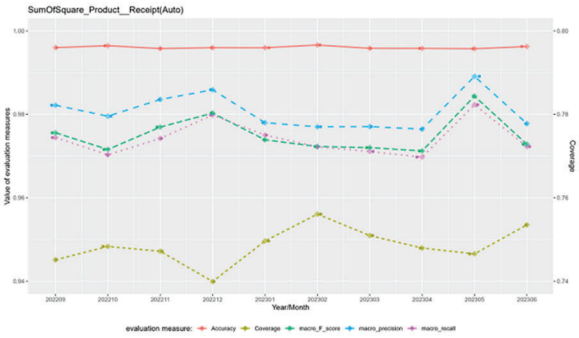
(b)信頼度：分割エントロピーに基づくハマーカー積



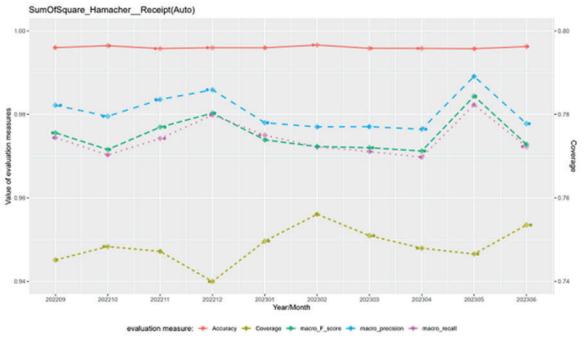
(c)信頼度：分割エントロピーに基づくアインシュタイン積



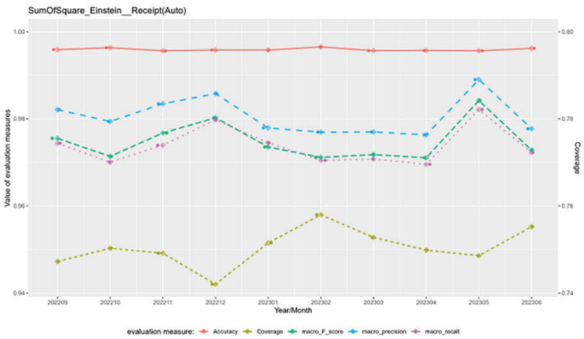
(d)信頼度：分割エントロピーに基づく最小値



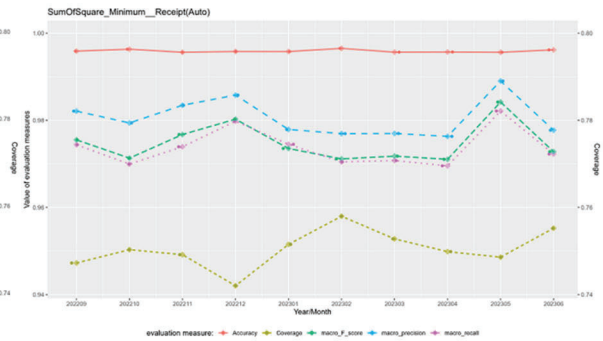
(e)信頼度：分割係数に基づく代数積



(f)信頼度：分割係数に基づくハマーカー積

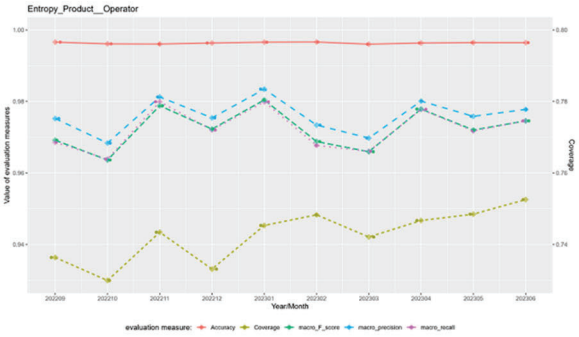


(g)信頼度：分割係数に基づくアインシュタイン積

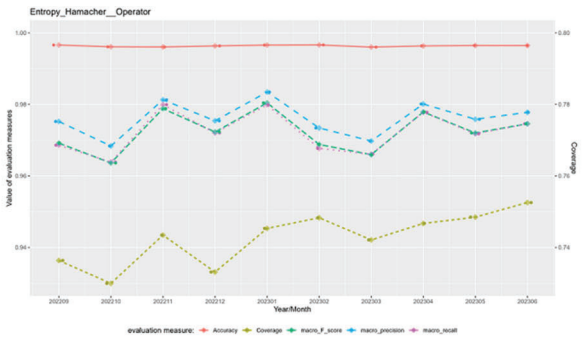


(h)信頼度：分割係数に基づく最小値

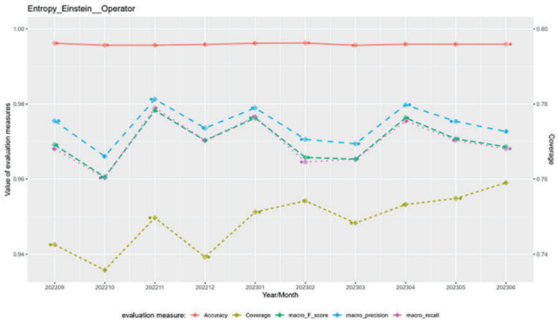
図 15. 文字認識したレシートデータにおける信頼度ごとの格付結果



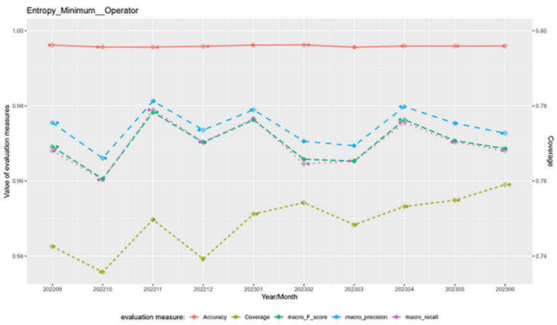
(a)信頼度：分割エントロピーに基づく代数積



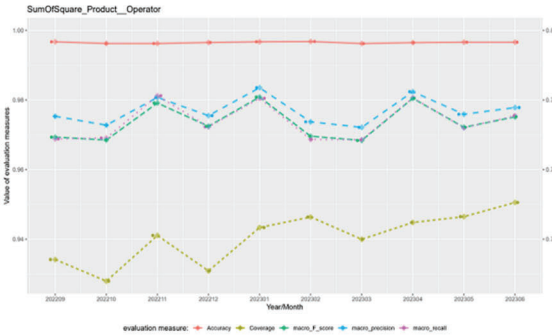
(b)信頼度：分割エントロピーに基づくハマーカー積



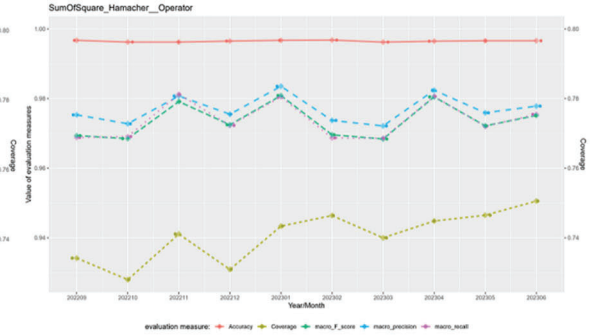
(c)信頼度：分割エントロピーに基づくアインシュタイン積



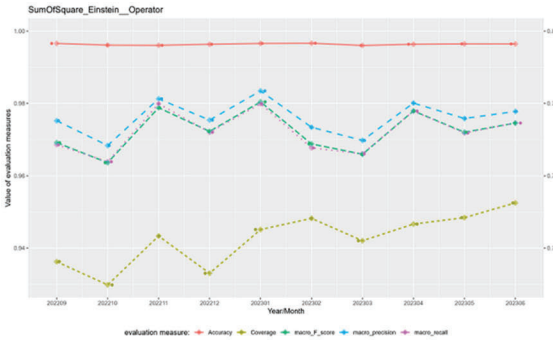
(d)信頼度：分割エントロピーに基づく最小値



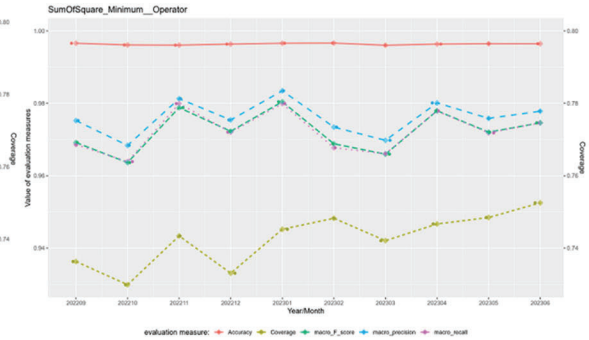
(e)信頼度：分割係数に基づく代数積



(f)信頼度：分割係数に基づくハマーカー積



(g)信頼度：分割係数に基づくアインシュタイン積



(h)信頼度：分割係数に基づく最小値

図 16. オペレータ入力したレシートおよび明細書のデータにおける信頼度ごとの格付結果

4．ファジィクラスタリングに基づくサポートベクターマシン

家計調査における自動格付のため、ファジィクラスタリングに基づくサポートベクターマシン(SVM)による新たな分類手法をいくつか提案した。SVM の確率スコアから得られる数値的特徴を用いた階層型の SVM に基づく分類手法はその一つである[29]。この手法では、ファジィ c-means 法(FCM)に基づく SVM による手法を実行し、クラスターごとに SVM を用いた符号の予測を行う。次に、クラスターごとに誤分類したデータを抽出し、誤分類データの各符号に対する SVM の確率スコアの平均および分散を符号ごとに計算し、各符号における誤りの傾向をとらえる。その後、新たな SVM 分類器を作成し、符号の再付与を行う。

本手法の数値例として、家計調査のオンラインデータを用いて性能評価を行った結果を示す。本検証では、収支項目分類の大分類符号に相当する分類に対して検証を行った。この分類には 17 の異なる符号が存在する。データは、2022 年 4 月及び 5 月の家計調査データから自動格付の対象となるデータとして抽出した約 180 万件のデータを用いた。このうちランダムに抽出された 90%（約 162 万件）を教師データとし、残りの 10%（約 18 万件）を評価データとして用いた。

表 1 に FCM に基づく SVM による各クラスターにおける符号ごとの分類結果を示す。また表 2 に FCM に基づく SVM による各クラスターにおける符号ごとの適合率（precision)、再現率（recall)、f1-スコア（f1-score）を示す。図 17、図 18 は、それぞれクラスター 1 とクラスター 2 の表 1 の正解率（accuracy）及び表 2 の適合率（precision)、再現率（recall)、f1-スコア（f1-score）を符号ごとに比較した結果を示したものである。

表 1 から、クラスター 2 の符号 1（カテゴリー：Food（食料))、符号 4（カテゴリー：Furniture & household utensils（家具・家事用品))、符号 9（カテゴリー：Culture & recreation(教養娯楽))、及び符号 10（カテゴリー：Other consumption expenditures（その他の消費支出））において、比較的多くの誤分類が起きていることがわかることから、これらの符号の誤分類データについて、その他の符号に対する SVM の確率スコアの平均および分散を符号ごとに計算した。図 19〜図 22 に、クラスター 2 の符号 1、符号 4、符号 9、及び符号 10 の誤分類データのその他の符号に対する SVM の確率スコアの平均及び分散を示す。この情報を用いて、クラスター 2 において、符号 1、符号 4、符号 9、符号 10 のデータについて、新たな SVM 分類器を作り、符号の再付与を行った。表 3 に FCM に基づく SVM（非階層型の SVM に基づく分類手法）と提案手法の階層型の SVM に基づく分類手法の分類結果の比較を示す。表 3 から符号の再付与の対象としたクラスター 2 の符号 1、符号 4、符号 9、符号 10 の全ての符号において正解率が向上したことがわかる。

表 1. 各クラスターにおける符号ごとの FCM に基づく SVM の分類結果

Label	Category	Cluster 1				Cluster 2			
		Number of text descriptions				Number of text descriptions			
		target data (a)	correctly predicted (b)	incorrectly predicted (a)-(b)	Accuracy (b)/(a)	target data (a)	correctly predicted (b)	incorrectly predicted (a)-(b)	Accuracy (b)/(a)
1	Food	54	45	9	0.833	119,328	118,965	**363**	0.997
2	Housing	38	30	8	0.789	292	201	91	0.688
3	Fuel, light & water charges	615	613	2	0.997	529	516	13	0.975
4	Furniture & household utensils	39	35	4	0.897	8,481	7,501	**980**	0.884
5	Clothing & footwear	2	2	0	1.000	2,563	2,348	215	0.916
6	Medical care	224	218	6	0.973	3,071	2,733	338	0.890
7	Transportation & communication	1,200	1,176	24	0.980	2,162	2,067	95	0.956
8	Education	123	106	17	0.862	276	238	38	0.862
9	Culture & recreation	724	702	22	0.970	4,917	4,122	**795**	0.838
10	Other consumption expenditures	904	849	55	0.939	4,635	3,837	**798**	0.828
11	Income	338	297	41	0.879	44	34	10	0.773
12	Receipts other than income	792	791	1	0.999	1	0	1	0.000
13	Non-consumption expenditures	2,637	2,631	6	0.998	37	26	11	0.703
14	Disbursements other than expenditures	2,206	2,150	56	0.975	1,165	979	186	0.840
15	Carry-over to next month	969	969	0	1.000	0	0	0	NA
16	Supplementary label	159	2	157	0.013	69	35	34	0.507
17	Consumption tax etc.	14,872	14,841	31	0.998	6,861	6,840	21	0.997
Total		25,896	25,457	439	**0.983**	154,431	150,442	3,989	**0.974**

表 2. 様々な評価関数による比較結果

Label	Category	Cluster 1			Cluster 2		
		precision	recall	f1-score	precision	recall	f1-score
1	Food	0.776	0.833	0.804	0.981	0.998	0.989
2	Housing	0.861	0.816	0.838	0.851	0.664	0.746
3	Fuel, light & water charges	0.997	0.997	0.997	0.970	0.974	0.972
4	Furniture & household utensils	0.833	0.897	0.864	0.932	0.886	0.908
5	Clothing & footwear	1.000	1.000	1.000	0.956	0.909	0.932
6	Medical care	0.982	0.969	0.975	0.957	0.878	0.916
7	Transportation & communication	0.912	0.980	0.945	0.956	0.953	0.955
8	Education	0.982	0.870	0.922	0.883	0.877	0.880
9	Culture & recreation	0.867	0.974	0.917	0.903	0.828	0.863
10	Other consumption expenditures	0.930	0.941	0.936	0.925	0.817	0.868
11	Income	0.914	0.879	0.896	0.829	0.773	0.800
12	Receipts other than income	1.000	0.999	0.999	0.000	0.000	0.000
13	Non-consumption expenditures	0.994	0.997	0.995	0.926	0.676	0.781
14	Disbursements other than expenditures	0.988	0.974	0.981	0.951	0.835	0.889
15	Carry-over to next month	1.000	1.000	1.000	NA	NA	NA
16	Supplementary label	0.500	0.006	0.012	0.854	0.507	0.636
17	Consumption tax etc.	0.997	0.998	0.997	0.998	0.996	0.997

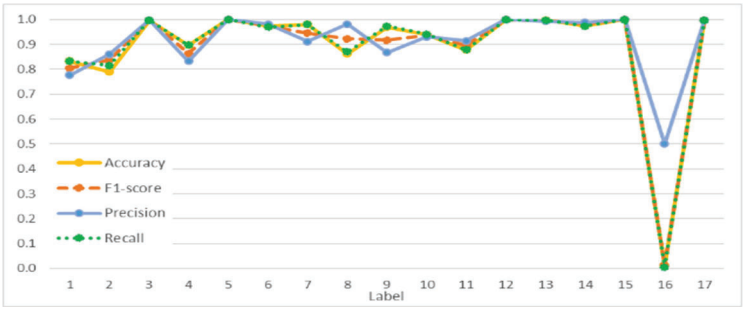


図 17. クラスター 1 における評価関数の値の比較

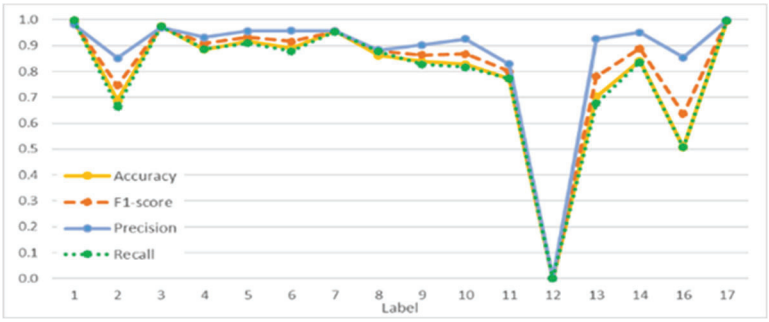
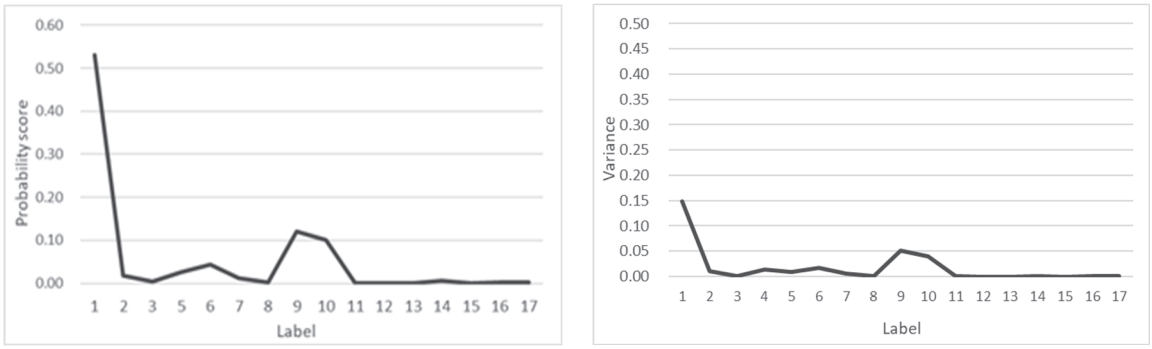
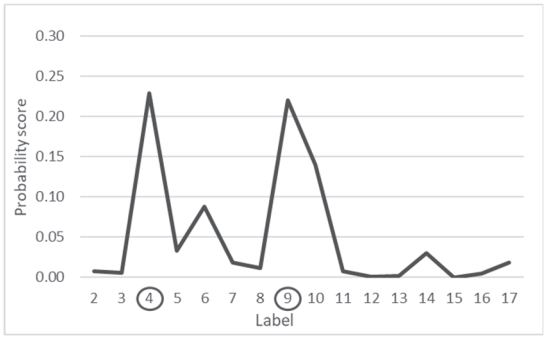


図 18. クラスター 2 における評価関数の値の比較

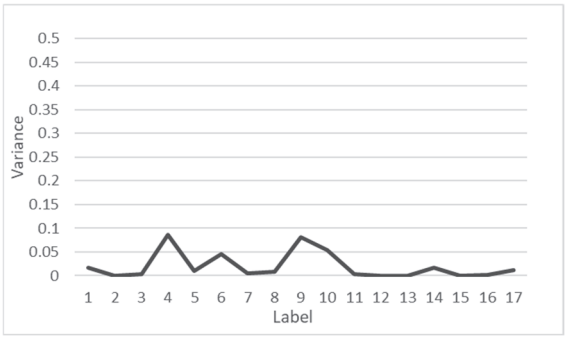


平均　　分散

図 19. クラスター 2 に含まれるデータのうち符号 1 から他の符号に誤格付された
データの確率スコアの平均と分散

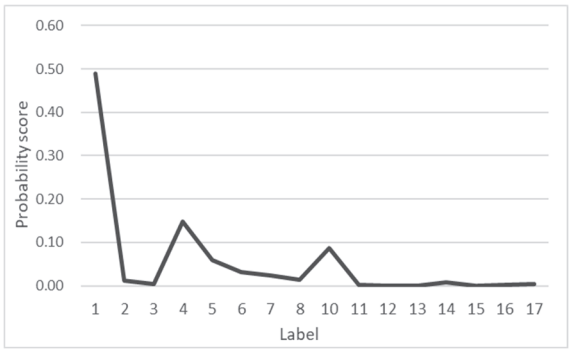


平均

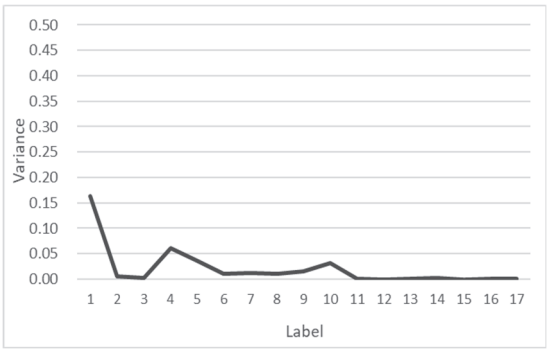


分散

図 20. クラスター 2 に含まれるデータのうち符号 4 から他の符号に誤格付された
データの確率スコアの平均と分散

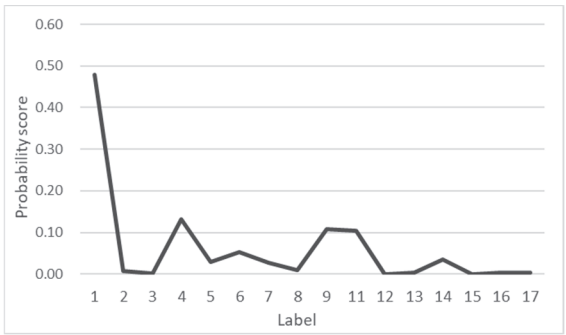


平均

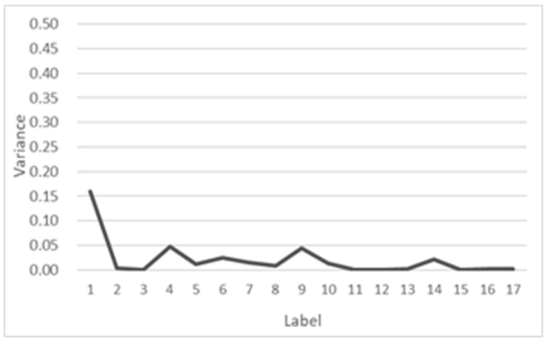


分散

図 21. クラスター 2 に含まれるデータのうち符号 9 から他の符号に誤格付された
データの確率スコアの平均と分散



平均



分散

図 22. クラスター 2 に含まれるデータのうち符号 10 から他の符号に誤格付された
データの確率スコアの平均と分散

表 3. 階層型の SVM に基づく分類手法と非階層型の SVM に基づく分類手法の格付結果の比較

Label	Category	Number of text descriptions of target data (a)	Original			Proposed		
			Number of text descriptions		Accuracy (b)/(a)	Number of text descriptions		Accuracy (b)/(a)
			correctly predicted (b)	incorrectly predicted (a)-(b)		correctly predicted (b)	incorrectly predicted (a)-(b)	
1	Food	119,328	118,965	363	0.997	119,009	319	0.997
4	Furniture & household utensils	8,481	7,501	980	**0.884**	7,848	633	**0.925**
9	Culture & recreation	4,917	4,122	795	0.838	4,275	642	0.869
10	Other consumption expenditures	4,635	3,837	798	**0.828**	4,159	476	**0.897**

さらに、強化学習を用いたファジィクラスタリングに基づく SVM による分類手法を二つ提案した [31]。 教師データのうちクリアなデータのみを SVM の教師データとして使用する手法と、全ての教師データを SVM の教師データとして使用する手法の二種の方法で数値実験を行った。家計調査のオンラインデータを用いて性能評価を行った結果を示す。本検証では、収支項目分類に対する分類精度について検証を行い、データには、2022 年 4 月及び 5 月の家計調査データから自動格付の対象となるデータとして抽出した約 180 万件のデータを用いた。このうちランダムに抽出された 90%（約 162 万件）を教師データとし、残りの 10%（約 18 万件）を評価データとして用いた。表 4 は、強化学習を用いたファジィクラスタリングに基づく SVM による分類結果である。表 4 の上半に全ての教師データを SVM の教師データとして使用した場合の結果、下半にクリアなデータのみを SVM の教師データとして使用した場合の結果を示している。それぞれの手法において、ファジィクラスタリングに基づく SVM を実行した場合、強化学習を用いたファジィクラスタリングに基づく SVM を実行した場合の分類結果の比較を行った。さらに、強化学習を用いたファジィクラスタリングに基づく SVM では、データにおける符号の出現頻度の降順に SVM をかけた場合（Fuzzy clustering based SVM with reinforcement learning (a)）とデータにおける符号の出現頻度の昇順に SVM をかけた場合（Fuzzy clustering based SVM with reinforcement learning (b)）の二つのパターンの予測方法を比較した。またそれぞれの方法において、現行のハイブリッド型格付支援システムの格付率と同程度となるように上から 76.5%のデータのみで評価した場合の結果と評価データ全てを評価した場合の結果について示している。また、表 5 は、提案手法からファジィクラスタリングの部分を除いた強化学習を用いた SVM による分類結果である。表 4 、表 5 から、全ての教師データを SVM の教師データとして使用し、データにおける符号の出現頻度の昇順で SVM をかけた場合の強化学習を用いたファジィクラスタリングに基づく SVM が他の手法に比べ優れていることがわかる。当手法では、上から 76.5%のデータに対して 0.9863 の正解率、また評価データ全体に対して 0.9578 の正解率で格付率 91.7%であった。さらに、適合率、再現率、f1-score の

評価指標においても当手法が他の手法に比べて良い結果であった。

表 6、表 7 は、表 4 、表 5 で示した各手法における SVM での符号付与にかかる計算時間を示している。表 6、表 7 より、全ての教師データを SVM の教師データとして使用した場合の強化学習を用いたファジィクラスタリングに基づく SVM (b) が最も計算時間が短くなることがわかる。この手法では、約 18 万件のデータに対して、約 164 秒で符号付与を行うことができたが、強化学習を用いない場合には約 18 万件のデータの符号付与に約 217 秒必要とした。

表 4 . 強化学習を用いたファジィクラスタリングに基づく SVM による分類結果

Method		Evaluation data	Coverage	Accuracy	Precision	Recall	F1 score
Use whole training dataset for training SVM	Fuzzy clustering based SVM	Evaluate top 76.5% of data	0.765	0.9860	0.9270	0.9150	0.9190
		Evaluate whole data	1.000	0.9178	0.8220	0.7730	0.7900
	Fuzzy clustering based SVM with reinforcement learning (a)	Evaluate top 76.5% of data	0.765	0.9862	0.9280	0.9150	0.9190
		Evaluate whole data	0.917	0.9575	0.8890	0.8700	0.8750
	Fuzzy clustering based SVM with reinforcement learning (b)	Evaluate top 76.5% of data	0.765	**0.9863**	0.9290	0.9180	0.9210
		Evaluate whole data	0.917	**0.9578**	0.8880	0.8730	0.8760
Use only clearer data for training SVM	Fuzzy clustering based SVM	Evaluate top 76.5% of data	0.765	0.9826	0.8990	0.8850	0.8880
		Evaluate whole data	1.000	0.9170	0.8070	0.7670	0.7790
	Fuzzy clustering based SVM with reinforcement learning (a)	Evaluate top 76.5% of data	0.765	0.9827	0.9010	0.8860	0.8890
		Evaluate whole data	0.911	0.9577	0.8750	0.8490	0.8560
	Fuzzy clustering based SVM with reinforcement learning (b)	Evaluate top 76.5% of data	0.765	0.9831	0.9080	0.8960	0.8980
		Evaluate whole data	0.911	0.9539	0.8710	0.8540	0.8550

表 5 . 強化学習を用いた SVM による分類結果

Method		Evaluation data	Coverage	Accuracy	Precision	Recall	F1 score
Use whole training dataset for training SVM	SVM	Evaluate top 76.5% of data	0.765	0.9857	0.9430	0.9330	0.9350
		Evaluate whole data	1.000	0.9173	0.8170	0.7620	0.7810
	SVM with reinforcement learning (a)	Evaluate top 76.5% of data	0.765	0.9858	0.9440	0.9330	0.9360
		Evaluate whole data	0.917	0.9572	0.8870	0.8620	0.8670
	SVM with reinforcement learning (b)	Evaluate top 76.5% of data	0.765	**0.9858**	0.9440	0.9350	0.9370
		Evaluate whole data	0.916	**0.9576**	0.8820	0.8640	0.8670
Use only clearer data for training SVM	SVM	Evaluate top 76.5% of data	0.765	0.9812	0.8930	0.8810	0.8810
		Evaluate whole data	1.000	0.9154	0.8030	0.7570	0.7720
	SVM with reinforcement learning (a)	Evaluate top 76.5% of data	0.765	0.9814	0.8940	0.8810	0.8820
		Evaluate whole data	0.912	0.9572	0.8590	0.8330	0.8380
	SVM with reinforcement learning (b)	Evaluate top 76.5% of data	0.765	0.9817	0.9020	0.8940	0.8910
		Evaluate whole data	0.911	0.9542	0.8480	0.8350	0.8340

表６．強化学習を用いたファジィクラスタリングに基づく SVM による処理時間

Method		Evaluation data	Processing time
Use whole training dataset for training SVM	Fuzzy clustering based SVM	Evaluate top 76.5% of data	217.11 sec.
		Evaluate whole data	
	Fuzzy clustering based SVM with reinforcement learning (a)	Evaluate top 76.5% of data	192.63 sec.
		Evaluate whole data	
	Fuzzy clustering based SVM with reinforcement learning (b)	Evaluate top 76.5% of data	**164.28 sec.**
		Evaluate whole data	
Use only clearer data for training SVM	Fuzzy clustering based SVM	Evaluate top 76.5% of data	213.09 sec.
		Evaluate whole data	
	Fuzzy clustering based SVM with reinforcement learning (a)	Evaluate top 76.5% of data	190.83 sec.
		Evaluate whole data	
	Fuzzy clustering based SVM with reinforcement learning (b)	Evaluate top 76.5% of data	168.02 sec
		Evaluate whole data	

表７．強化学習を用いた SVM による処理時間

Method		Evaluation data	Processing time
Use whole training dataset for training SVM	SVM	Evaluate top 76.5% of data	213.99 sec.
		Evaluate whole data	
	SVM with reinforcement learning (a)	Evaluate top 76.5% of data	186.41 sec.
		Evaluate whole data	
	SVM with reinforcement learning (b)	Evaluate top 76.5% of data	168.02 sec.
		Evaluate whole data	
Use only clearer data for training SVM	SVM	Evaluate top 76.5% of data	208.95 sec.
		Evaluate whole data	
	SVM with reinforcement learning (a)	Evaluate top 76.5% of data	188.10 sec.
		Evaluate whole data	
	SVM with reinforcement learning (b)	Evaluate top 76.5% of data	169.29 sec.
		Evaluate whole data	

５．まとめ

本稿では、機械学習型格付支援について、様々な手法を紹介すると共に、数値実験よりこれらの提案手法の有効性を示した。これらの手法の詳細については、参考文献を参照されたい。また、これらの手法を正確に用いるためには、本稿で示した評価指標のみならず、T-norm や信頼度の違いによる数学的性質の差異に応じた適切な検証が必要であり、それを行わず、単に、ここで述べた指標のみを用いて、T-norm の種類、信頼度の関数の種類、ファジィクラスタリング手法、サポートベクターマシンの種類を決定す

ることは、作成されるデータの品質の劣化を招くことになり得ることを十分に理解した上で、用いる必要があることに注意されたい。

参考文献

[1] Benedikt, Lanthao, Chaitanya Joshi, Louisa Nolan, Nick De Wolf, Barry Schouten. (2020), “Optical character recognition and machine learning classification of shopping receipts”, Available at: https://ec.europa.eu/eurostat/documents/54431/11489222/6+Receipt+scan+analysis.pdf

[2] Bengio, Yoshua, Rejean Ducharme, Pascal Vincent, Christian Jauvin. (2003), “A neural probabilistic language model”, Journal of Machine Learning Research **3**: 1137-1155

[3] Bezdek, James C. (1981), “Pattern Recognition with Fuzzy Objective Function Algorithms”, Plenum Press, New York

[4] Bezdek, James C., James Keller, Raghu Krisnapuram, Nikhil R. Pal. (1999), “Fuzzy Models and Algorithms for Pattern Recognition and Image Processing”, Kluwer Academic Publishers

[5] Brown, Tom B., Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, Sandhini Agarwal, Ariel Herbert-Voss, Gretchen Krueger, Tom Henighan, Rewon Child, Aditya Ramesh, Daniel M. Ziegler, Jeffrey Wu, Clemens Winter, Christopher Hesse, Mark Chen, Eric Sigler, Mateusz Litwin, Scott Gray, Benjamin Chess, Jack Clark, Christopher Berner, Sam McCandlish, Alec Radford, Ilya Sutskever, Dario Amodei. (2020), “Language Models are Few-Shot Learners”, arXiv:2005.14165

[6] Cristianini, Nello, John Shawe-Taylor. (2000), “An Introduction to Support Vector Machines and other kernel-based learning methods”, Cambridge University Press

[7] Guidotti, Riccardo, Anna Monreale, Salvatore Ruggieri, Dino Pedreschi, Franco Turini, Fosca Giannotti. (2018), "Local rule-based explanations of black box decision systems" arXiv preprint arXiv:1805.10820

[8] Gweon, Hyukjun, Matthias Schonlau, Lars Kaczmirek, Michael Blohm, and Stefan Steiner. (2017), "Three methods for occupation coding based on statistical learning", Journal of Official Statistics, 33 (1): 101-122.

[9] Hacking, Wim, Leon Willenborg (2012). "Coding; interpreting short descriptions using a classification", Statistics Methods, Statistics Netherlands, Available at: https://www.cbs.nl/en-gb/our-services/methods/statistical-methods/throughput/throughput/coding.

[10] Jentoft, Susie, Toth Boriska, Müller Daniel. (2022) “From Manual to Machine: challenges in machine learning for COICOP coding”, in the 29th Nordic Statistical Meeting (NSM2022)

[11] Ito, Haruto. (2023) “Comparison of Results between SVM and FCM - For Comparison of Supervised and Unsupervised Classification -”, College of Social Engineering, University of Tsukuba Graduation Thesis (Supervisor Prof. Dr. M. Sato-Ilic) (in Japanese)

[12] Kenton, Zachary, Tom Everitt, Laura Weidinger, Iason Gabriel, Vladimir Mikulik, Geoffrey Irving. (2021), “Alignment of Language Agents”, DeepMind, arXiv:2103.14659

[13] Kudo, Taku, Kaoru Yamamoto, Yuji Matsumoto. (2004), “Applying conditional random fields to Japanese morphological analysis”, in the 2004 Conference on Empirical Methods in Natural Language Processing, Barcelona, Spain, 25-26, Jul. 2004, 230-237

[14] Martindale, Hazel, Edward Rowland, Tanya Flower, Gareth Clews. (2020) “Semi-supervised machine learning with word embedding for classification in price statistics”, Data & Policy, **2**, 2020, e12, Cambridge University Press

[15] Menger, Karl. (1942), "Statistical metrics", in Proceedings of the National Academy of Sciences of the United States of America, **28**: 535-537

[16] Mikolov, Tomas, Kai Chen, Greg S. Corrado, Jeffrey Dean. (2013), “Efficient estimation of word representations in vector space”, arXiv preprint arXiv:1301.3781

[17] Mizumoto, Masaharu. (1989), "Pictorical representation of fuzzy connectives, Part I: Cases of T-norms, t-Conorms and Averaging Operators", Fuzzy Sets and Systems, **31**: 217-242

[18] Platt, John C. (1999), “Probabilistic Outputs for Support Vector Machines and Comparisons to Regularized Likelihood Methods”, In Advances in Large Margin Classifiers, Alexander J Smola, Peter Bartlett, Bernhard Scholkopf, and Dale Schuurmans (eds) MIT Press

[19] Ribeiro, Marco Tulio, Sameer Singh, Carlos Guestrin. (2018), “Anchors: High-precision model-agnostic Explanations", The Thirty-Second AAAI Conference on Artificial Intelligence: 1527-1535

[20] Solaiman, Irene, Christy Dennison. (2021), “Process for Adapting Language Models to Society (PALMS) with Values-Targeted Datasets”, 35th Conference on Neural Information Processing Systems (NeurIPS 2021), arXiv:2106.10328

[21] Schweizer, Berthold, Abe Sklar. (2005), Probabilistic metric spaces, Dover Publications

[22] Takaoka, Kazuma, Sorami Hisamoto, Noriko Kawahara, Miho Sakamoto, Yoshitaka Uchida, Yuji Matsumoto. (2018), “Sudachi: a Japanese Tokenizer for Business”, Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC): 2246-2249, European Language Resources Association

[23] Toko, Yukako, Kazumi Wada, Shinya Iijima, and Mika Sato-Ilic. (2018a) “Supervised Multiclass Classifier for Autocoding Based on Partition Coefficient”, Ireneusz Czarnowski, Robert J. Howlett, Lakhmi C. Jain, and Ljubo Vlacic (eds), Smart Innovation, Systems and Technologies 97: 54-64. Springer, Switzerland (Best research paper award 受賞)

[24] Toko, Yukako, Shinya Iijima, Mika Sato-Ilic. (2018b), "Overlapping classification for autocoding system", Journal of Romanian Statistical Review, **4**: 58-73

[25] Toko, Yukako, Shinya Iijima, Mika Sato-Ilic. (2019),” Generalization for improvement of the reliability score for autocoding”, Journal of Romanian Statistical Review, **3**: 47-59

[26] Toko, Yukako, Mika Sato-Ilic. (2020), “Improvement of the training dataset for supervised multiclass classification” , Czarnowski, I., Howlett, R.J., Jain, L. C. (Eds.), Intelligent Decision Technologies, Smart Innovation, Systems and Technologies, Springer, Singapore, **193**: 291-302

[27] Toko, Yukako, Mika Sato-Ilic. (2022a), “Autocoding based Multiclass Support Vector Machine by Fuzzy c-Means”, Romanian Statistical Review, **1**: 27–39

[28] Toko, Yukako, Mika Sato-Ilic. (2022b), “Fuzzy Clustering based Support Vector Machine for Autocoding and Interpretation of Results”, Book of Abstract of IASC-ARS 2022 Interim conference: 43-44

[29] Toko, Yukako, Mika Sato-Ilic. (2023), “Hierarchical Support Vector Machine based Classifier for Autocoding”, G. A. Tsihrintzis, C. Toro, S. A. Rios, R. J. Howlett, L. C. Jain (Eds.) Procedia Computer Science, **225**: 2733-2742

[30] Toko, Yukako, Mika Sato-Ilic. (2024a), “Difference on Evaluation Scores Considering Image Descriptions for Autocoding”, Romanian Statistical Review, **1**:27-42

[31] Toko, Yukako, Mika Sato-Ilic. (2024b), “Comparison of fuzzy clustering based SVM with reinforcement learning based SVM for autocoding of the Family Income and Expenditure Survey”, C. Toro, S. A. Rios, R. J. Howlett, L.C. Jain (Eds.) Procedia Computer Science, **246**: 1820-1829

[32] Statistics Bureau of Japan. “Outline of the Family Income and Expenditure Survey”, Available at: https://www.stat.go.jp/english/data/kakei/1560.html