社会学理論と言語処理技術の接続：ブルデュー理論に基づく言語処理「界」の分析を事例に

高橋祐貴[1] 塚越柚季[2]

[1]フリー [2]東京大学大学院人文社会系研究科
biz.yuki.takahashi@gmail.com yuzuki@l.u-tokyo.ac.jp

概要

本研究は，人文学の系譜に位置付く社会学理論の実証の文脈において，いかに自然言語処理が応用可能かを示すべく，言語処理の研究者集団を対象にブルデュー流の「界」の分析を行い，その分化の構造を明らかにする．ACL Anthology と researchmap を用いて構築した有力な言語処理研究者 578 名のデータセットを用いて，多重対応分析により「界」の空間を構築したところ，資本総量の差は見られたものの，解釈可能な資本構成の交差配列構造は認められなかった．BERTopic により析出した研究分野との関連もそれほど明確とは言えず，理工系の分野においてはブルデュー理論が有効とは言えないことが示唆された．

1 はじめに

哲学の流れを汲む社会理論に端を発しながらも，実証科学としての側面を強調する学問として発展してきた社会学では，言葉による定式化されていない理論を，定量・定性それぞれの分析により，経験的に検証する試みが積み重ねられてきた．その中で，特に定量的な実証分析において近年関心が高まっているのが，自然言語処理の技術である．その範囲は，LDA を始めとするトピックモデルの応用[1,2]から，単語埋め込みベクトルによる単語の意味変遷の推定[3]，LLM を用いた政治的反応の予測実験[4]など，非常に幅広い．

社会学が自然言語処理を応用するに当たって，特に強く意識されるのは，分析により得られた何らかの数値的結果が社会学理論上ではどのような意味を持つのか，その対応を明確化することである．例えば，LDA により得られるトピックの理論的な位置付けを，Alfred Schutz のレリヴァンス理論を参照し，「解釈的レリヴァンスの計量化」として位置付ける，といった具合である[1]．本研究は，このような社会学における自然言語処理の応用の一例として，2 節にてフランスの社会学者 Pierre Bourdieu の「界」理論における「態度決定」という概念とトピックモデルを対応させた分析枠組みを提示する．そして 3 節で実際の分析例として，日本の言語処理「界」の分析を行うためのデータと方法を提示し，4 節にて人文社会系の分野の分析手法として提案されたブルデューの学問の「界」の分析視角が，理工系分野である言語処理の分析でも有効性を持つかを検証する．5 節では，得られた知見を基に，今後の研究の展開を議論する．

2 背景と関連研究

2.1 ブルデューの「界」理論と学問分析

「文化資本」の概念などで知られるブルデューの「界」理論は，元々は作品の創作者の社会関係と関連づけて作品の創作と受容を考察するべく，芸術分析の流れの中で提示された．まず「資本」の変数を用いて多重対応分析を行い，2〜3 次元の「界」の空間を析出する．「資本」は，例えば文芸界なら文学賞の受賞歴や部数，科学界なら論文数や学会賞歴など，それぞれの「界」の中で有利な立場に立つために必要な有形・無形の所有物が分析ごとに定められる．続いてこの空間における座標を行為者の「位置」とし，座標軸の意味を解釈することで「界」の構造を読み解く．多くの場合，最も分散を説明する 1 軸目に資本の多寡が反映され，2 軸目に資本の構成割合の分化が見られるという解釈になる．そして，「界」における「位置」と，行為者が表明する創作物の内容や政治的態度といった「態度（位置）決定」は関連する，という図式を通じて，文化的生産物と行為者の社会的位置の関係性を読み解くのが，ブルデューの「界」理論の枠組みである[5]．

ブルデュー自身は『ホモ・アカデミクス』[6]において，フランスの人文社会科学界の分析を行っている．ここで，ブルデューは多様な学問的所有物を基準に選別した，界の中で高い位置を占めると想定されるパリの「文学系」正教授のサンプルから，対応分析により文学・人間科学部の権力空間を構築した．ブルデューが見出したのは，「純粋に大学的な権力」と，「多少とも制度化された知的名声」（学問的威信と知識人界での名声という象徴資本の双方を含む）の 2 種類の権力の対立である．加えてもう一つ主要に見出された区分として，単に大学的・学問的権威の多い年長の教授と，大学内権力としては下位に位置付く若年の教授たちという区分があるとした．この年齢的な区分は，資本総量の分化に対応するものとして捉えられる．こうしてブルデューは，文学・人間科学部界の中にも，相対的に文化資本を多く持つグループと，経済資本を多く持つグループがあるという対立構造を見出した．

さらにブルデューが踏み込んで検討したのは，文学・人間科学部界における知的生産物の内容や形式と，こうした権力空間との間に見られる，構造的な関連性についてである．文学テクストの正統的注釈を主張した保守派のレイモン・ピカールと現代主義的注釈を主導した革新派のロラン・バルトが，対応分析の図では対照的な位置にいることが，その一例であった．

2.2 経験的な先行研究

後続の学問の界の研究では，フランスの経済学界[7]，アメリカの社会学界[8]などの分析が行われ，比較的明瞭な資本構成の対立構造が見られること，また界における「位置」と研究分野や方法といった「態度決定」とが関連することが示唆されてきた．特に「態度決定」の分析は，ブルデュー自身は定性的に読解した著作物の内容と，対応分析の結果とを照らし合わせて行っていたのに対し，後続の研究は軸の生成に寄与しないサプリメンタリ変数として内容に対応する変数（例えば研究者個々人の研究分野）を投影する方法へと変化していった．ここで，「態度決定」が「創作物の空間」であるとしたブルデューの主張に立ち返り，トピックモデルを用いて論文や書籍の本文から「態度決定」変数を導き出す手法を取ったのが，日本の社会学界を分析した高橋[9]である．ここでは論文誌の本文データに構造トピックモデルを適用し，析出したトピック割合を基に個々の研究者の主要な研究分野と方法の変数を作成している．ブルデューが定性的に行なっていた分析をより大規模に拡張する上で，トピックモデルにより析出したトピックを「態度決定」として見なす手法には応用可能性がある．

一方で後続の先行研究が欠いてきたのが，いわゆる「理系」の学問分野の「界」の分析である．元々芸術創作との類似性から人文社会系の学問に多く適用されてきた「界」理論が，理工系や医薬系といった別の学問分野においても有効性を持つかは検証されてこなかった．そこで本研究は，理工系の一分野と見做せる自然言語処理を対象に，「界」の構造の分析を行い，人文社会系分野と同様の資本構成の対立構造や「位置」と「態度決定」との対応関係が見られるかを検証する．

3 データと方法

3.1 データ

本研究で用いるのは，ACL Anthology と researchmap を基に構築された，有力な言語処理研究者 578 名の「資本」と「態度決定」に関するデータセットである．まず，ACL Anthology より日本人と判別できる氏名に紐づく発表実績を抽出し，その氏名や発表物のタイトルから researchmap の個人ページを特定し，合計 1443 名の研究者の研究実績や役員歴，受賞歴などのデータを収集した．そこからさらに researchmap で「論文」の項目に登録のある研究者にデータを絞り込んだ結果，サンプルサイズは 578 となった．構築されたサンプル集団は，ACL Anthology に何らかの論文を掲載しておりかつ researchmap に業績を管理しているという意味でバイアスがかかっている点には注意されたい．変数構築の詳細は付録に記載している．

3.2 方法

まず，「資本」の変数を用いて界の空間を構築し，各変数の平均的な配置から軸の意味を読み解く．続いて「態度決定」変数として研究分野の変数をサプリメンタリ変数として投影し，研究分野と「界」の構造に何らかの関連が見られるかを分析する．

研究分野変数は，ACL Anthology のアブストラクトのデータに対し BERTopic[10]を適用し，析出したトピックの割合が多い順に 2 分野を，その研究者の主要研究分野と見做して作成した．析出したトピッ

クとその代表語は表 1 の通りである．

表 1　BERTopic によるトピックの解釈

解釈	Representation
言語処理	['translation', 'language', 'models', 'model', 'based', 'method', 'paper', 'data', 'task', 'results']
機械翻訳	['translation', 'machine', 'english', 'nmt', 'parallel', 'japanese', 'sentence', 'corpus', 'sentences', 'paper']
埋め込みモデル	['models', 'word', 'embeddings', 'tasks', 'language', 'model', 'bias', 'method', 'performance', 'embedding']
対話	['dialogue', 'responses', 'user', 'utterances', 'systems', 'human', 'response', 'users', 'task', 'dialog']
日本語処理	['japanese', 'corpus', 'patent', 'simplification', 'word', 'words', 'method', 'english', 'functional', 'expressions']
構文解析	['parsing', 'argument', 'dependency', 'structure', 'event', 'resolution', 'coreference', 'structures', 'discourse', 'parser']
固有表現認識、関係認識	['entity', 'entities', 'relation', 'ner', 'extraction', 'knowledge', 'model', 'relations', 'embeddings', 'task']
要約生成	['summarization', 'summaries', 'generation', 'headline', 'summary', 'document', 'model', 'text', 'information', 'data']
音声認識	['speech', 'recognition', 'asr', 'acoustic', 'language', 'translation', 'sign', 'censrec', 'spoken', 'japanese']
SNS分析	'twitter', 'polarity', '19']
マルチモーダル（画像、動画）	['image', 'video', 'images', 'visual', 'captions', 'multimodal', 'caption', 'videos', 'captioning', 'dataset']
機械翻訳データセット（辞書）	['mt', 'dictionaries', 'translation', 'lrs', 'resources', 'web', 'www', 'machine', 'user', 'language']
質問応答	['question', 'answer', 'questions', 'answering', 'qa', 'rc', 'model', 'reading', 'mrc', 'scoring']
文法修正	['gec', 'correction', 'grammatical', 'error', 'evaluation', 'learners', 'models', 'examples', 'learner', 'manual']
AI人狼	['game', 'agents', 'agent', 'werewolf', 'games', 'llms', 'logics', 'contest', 'circumstances', 'holding']
議事録処理	['minutes', 'assembly', 'local', 'retrieval', 'spoken', 'meeting', 'corpus', 'document', 'task', 'queries']
音声感情分析	['emotion', 'emotional', 'speech', 'dialogue', 'labeling', 'fusion', 'labels', 'modalities', 'corpus', 'learning']
フィードバック生成	'ranking', 'recommendations', 'generation', 'recommendation', 'online', 'task']
金融	['stock', 'prices', 'market', 'price', 'comments', 'numerical', 'weather', 'changes', 'data', 'generating']
実況生成	['commentaries', 'commentary', 'game', 'live', 'sports', 'races', 'domain', 'real', 'world', 'spectators']
LLM	['reasoning', 'logical', 'syllogistic', 'llms', 'cot', 'biases', 'models', 'abilities', 'symbolic', 'step']
トピックモデル	['topic', 'model', 'topics', 'adverbs', 'tree', 'dynamic', 'hierarchical', 'words', 'structured', 'interactive']
物語生成	['story', 'event', 'stories', 'events', 'module', 'sequence', 'generation', 'narrative', 'salience', 'schema']
広告	'appeals', 'concerns']
料理レシピ	'procedural', 'food', 'dish']
数値処理、Numerical Commonsense	['numerals', 'numeral', 'numbers', 'digit', 'ncs', 'quantitative', 'nn', 'arithmetic', 'magnitudes', 'models']
自動運転	['driving', 'channel', 'dialog', 'utterances', 'driver', 'rate', 'utterance', 'human', 'car', 'context']
フランス語（その他言語）	['des', 'les', 'la', 'le', 'une', 'et', 'du', 'dans', 'pour', 'en']

4 分析結果

「資本」の変数を用いて多重対応分析を行なった結果，1 軸の説明率が高く，修正分散説明率は 2 軸で 8 割を超えた（表 2）．そのため，2 軸までの結果を主に解釈する．

表 2　多重対応分析の分散説明率

各軸の分散説明率

次元	固有値	分散説明率	ベンゼクリ補正比率	累積ベンゼクリ補正比率
dim 1	0.19	13.13	77.83	77.83
dim 2	0.07	5.22	8.89	86.72
dim 3	0.06	4.00	4.33	91.04
dim 4	0.05	3.20	2.23	93.27
dim 5	0.04	3.00	1.81	95.08

4.1 軸の解釈

1 軸への寄与率が平均以上のモーダルを 2-1 軸上

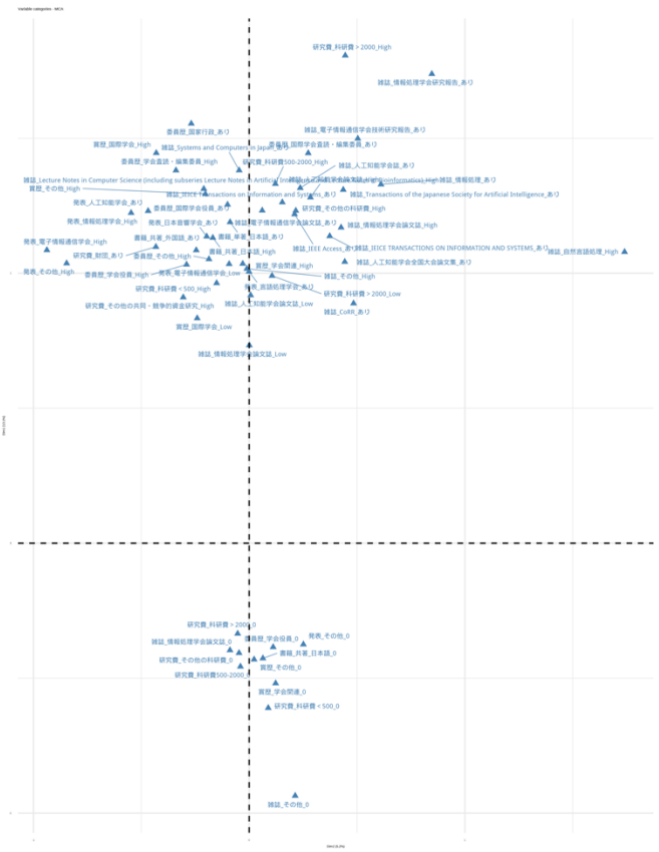

図 1　1 軸への寄与率が高いモーダル(2-1 軸)

の空間に投影したものが図 1 である．

これを見ると，1 軸方向には資本総量と解釈できそうな分化が反映されていることが分かる．1 軸方向に正の座標には雑誌論文や学会発表，研究費の獲得が高いモーダルが並び，負の方向には 0 のモーダルが並ぶ．

他方，2 軸の構成に平均以上に寄与しているモーダルをプロットしたのが図 2 である．これを見ると，まず寄与率が高い変数が少ない上に，査読のない一部の学会発表が多い群と，査読付きの一部の雑誌論文や海外学会での発表が多い群の分化を示している，程度にしか解釈できず，資本構成の明確な分化が認められるとは言い難いことが分かる．

4.2 分野変数の投影と解釈

2-1 軸の「界」の空間上に，「態度決定」としての研究分野変数をサプリメンタリ変数として投影したのが図 3 である．多くの分野が原点近くに配置され，

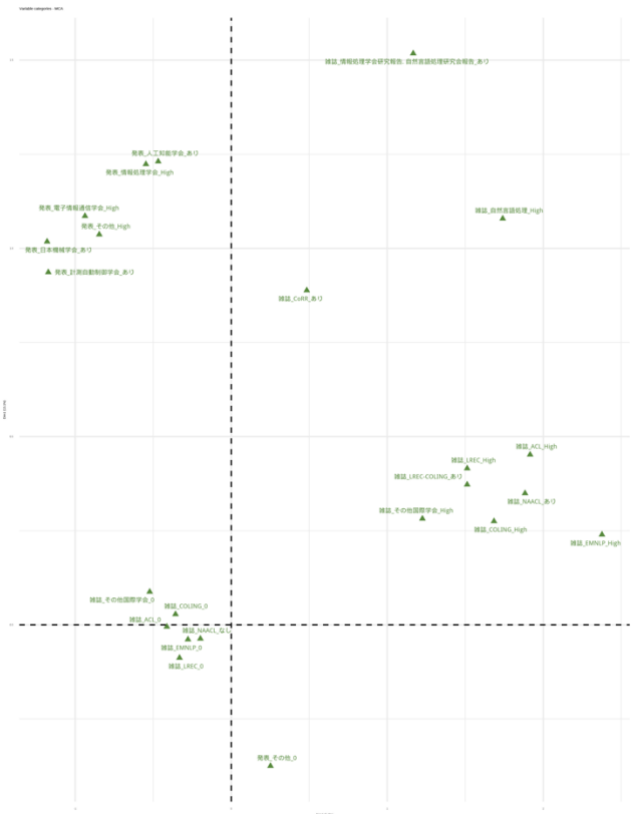

図 2　2 軸への寄与率が高いモーダル(2-1 軸)

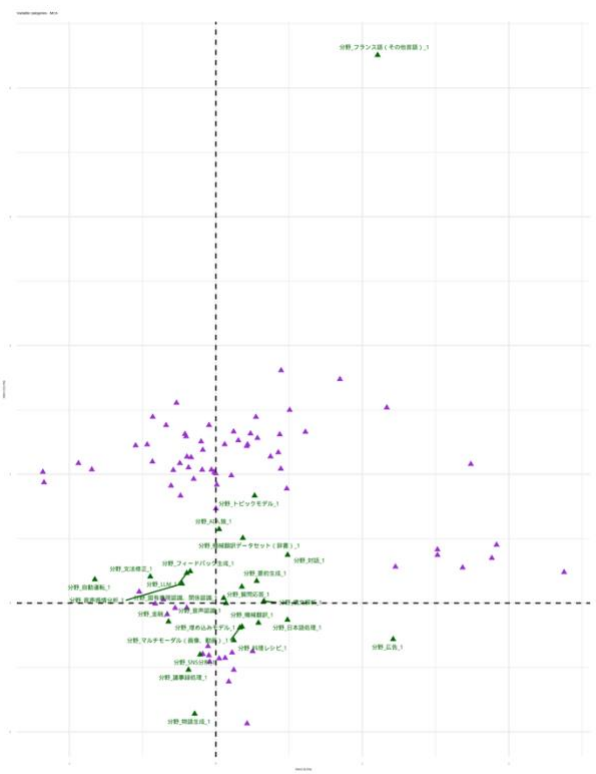

図 3　分野変数の投影(2-1 軸)

界における位置と態度決定の対応があると言い難い．

5 おわりに

本研究では，社会学における自然言語処理の応用例を示すべく，ブルデュー理論に基づいた言語処理「界」の分析を試みた．結果として，言語処理分野においては，ブルデュー理論と整合的な界の分化は確認できないことが明らかになった．同様に researchmap のデータを基に社会学者に関する分析を行なった高橋[9]では解釈可能性の高い資本構成の分化が見られたのとは対照的である．これは，人文社会系の分野に見られた内部対立構造が言語処理分野には見られない可能性を示唆する．一方でこの結果自体は，人文社会系分野で想定されるような「資本」変数では分化の構造を捉えられないからとも考えられるし，researchmap をデータソースとして用いることでそもそも分野の全体を捕捉できなかったという問題も考えられる．科研費申請に用いられる researchmap は企業所属の研究者を捕捉できない問題が指摘されている[11]が，多くを大学所属の研究者が占める人文社会系の分野と異なり，企業所属の研究者も中心的な役割を果たしている言語処理においてこれは大きな欠点と言えよう．また，個人が主体の基礎単位となる人文社会系とは異なり，理工系分野では研究室単位で研究主体を捉えた方が適切だとも考えられるだろう．今後ブルデュー派の分析を理工系分野に展開していくに当たっては，人文社会系とのこうした性質の違いを考慮し，より適切なデータソースを検討していく必要がある．

また今後の展開としては，その他の自然言語処理技術とブルデューの理論との接合を模索する道筋も考えられる．Mochihashi [12]が提案した Researcher2Vec や，研究タイトルから研究者の埋め込みを得た Nagao & Katsurai[13]のように「書かれたものの内容から研究者の相対的な位置を把握する」技術はすでに複数存在している．例えば「態度決定」の空間を Researcher2Vec と通じて構築し，その上で関心のあるその他の変数の平均的な配置が分化するかを見る手もあるだろう．自然言語処理の社会学への応用はまだ始まったばかりであり，今後も発展が期待される．

参考文献

1. **小田中悠・中井豊,** 2019, 「意味世界の計算社会科学的分析に向けて――社会学におけるトピックモデルの意義の検討」『理論と方法』34(2): 280‐95.

2. **瀧川裕貴,** 2019, 「戦後日本社会学のトピックダイナミクス」『理論と方法』34(2): 238‐61.

3. **Kozlowski, Austin C., Matt Taddy & James A. Evans,** 2019, "The Geometry of Culture: Analyzing the Meanings of Class through Word Embeddings," *American Sociological Review*, 84(5): 905–49.

4. **Kozlowski, Austin, Hyunku Kwon & James Evans,** 2024, "In Silico Sociology: Forecasting COVID-19 Polarization with Large Language Models."

5. **磯直樹,** 2020, 『認識と反省性――ピエール・ブルデューの社会学的思考』法政大学出版局.

6. **Bourdieu, Pierre,** 1984, *Homo academicus*, MINUIT. （石崎晴己・東松秀雄訳, 1997, 『ホモ・アカデミクス』藤原書店.）

7. **Lebaron, Frédéric,** 2001, "ECONOMISTS AND THE ECONOMIC ORDER The Field of Economists and the Field of Power in France," *European Societies*, 3(1): 91–110.

8. **Warczok, Tomasz & Stephanie Beyer,** 2021, "The Logic of Knowledge Production: Power Structures and Symbolic Divisions in the Elite Field of American Sociology," *Poetics*, 87: 101531.

9. **高橋祐貴,** 2023, 「現代日本の社会学界――ブルデュー派アプローチによる『界』の構造と学知の関連の分析」東京大学大学院教育学研究科 2023 年度修士論文.

10. **Grootendorst, Maarten,** 2022, "BERTopic: Neural Topic Modeling with a Class-Based TF-IDF Procedure," *arXiv preprint arXiv:2203.05794*.

11. **渡辺健太郎,** 2023, 「日本の研究者を対象とした無作為抽出調査は可能か」科学技術社会論学会第 22 回年次研究大会発表資料.

12. **Mochihashi, Daichi,** 2023, "Researcher2Vec: Neural Linear Model of Scholar Recommendation for Funding Agencies," International Society for Scientometrics and Informatics (ISSI 2023), 2: 329-335.

13. **Nagao, Hiroyoshi & Marie Katsurai,** 2024, "Researcher Representations Based on Aggregating Embeddings of Publication Titles: A Case Study in a Japanese Academic Database," Proceedings of the Fourth Workshop on Scholarly Document Processing (SDP 2024): 277–82.

A 付録　変数の構築方法について

対応分析に用いた変数は以下の要領で作成した．まず，ACL Anthology に収録されている国際学会報告については主要会議（LREC，ACL，EMNLP，NAACL，COLING，LREC-COLING）での発表を会議名を用いて抽出し，これらでの発表数をカウントした．それ以外の ACL Anthology への掲載は全て「その他国際学会」変数でカウントした．国際学会報告以外の「資本」変数は，すべて researchmap の当該項目のデータからカウントした．登録のない欠損データについては全て 0 件と見做して欠損を補充している．分野の変数は，ACL Anthology の Abstract の項目にトピック数を 28 と設定して BERTopic を適用し，余剰カテゴリである「-1」に分類されたもの以外の 27 トピックを解釈して作成した．トピック割合を計算するオプションを用いて各アブストラクトごとにトピック割合を計算し，それを個人名で集計して割合が多い上位 2 トピックを，その研究者の研究トピックと見做して紐づけた．下記の表に各変数の度数分布と対応分析における座標を示す．

表　アクティブ変数とサプリメンタリ変数の度数分布と座標

variable	value	頻度	割合	対応分析における座標 Dim 1	Dim 2
アクティブ変数					
委員歴_国家行政	なし	555	96	−0.06442	0.011119
	あり	23	4	1.554509	−0.2683
委員歴_国際学会役員	なし	526	91	−0.12399	0.009905
	あり	52	9	1.254221	−0.10019
委員歴_国際学会査読・編集委員	なし	542	93.8	−0.09598	−0.01819
	あり	36	6.2	1.444996	0.273825
委員歴_地方行政	なし	565	97.8	−0.01954	0.016904
	あり	13	2.2	0.849082	−0.73469
委員歴_学会役員	0	368	63.7	−0.38495	0.111252
	Low	123	21.3	0.421679	−0.12768
	High	87	15.1	1.032142	−0.29007
委員歴_学会査読・編集委員	0	449	77.7	−0.25527	0.041378
	Low	81	14	0.596588	−0.02834
	High	48	8.3	1.381114	−0.33922
委員歴_その他	0	411	71.1	−0.31942	0.103837
	Low	86	14.9	0.535368	−0.32081
	High	81	14	1.052336	−0.18626
書籍_共著_外国語	なし	485	83.9	−0.20814	0.047033
	あり	93	16.1	1.08544	−0.24528
書籍_共著_日本語	0	384	66.4	−0.42648	0.063617
	Low	97	16.8	0.556888	−0.08341
	High	97	16.8	1.13146	−0.16844
書籍_単著_外国語	なし	556	96.2	−0.05173	0.0207
	あり	22	3.8	1.307438	−0.52313
書籍_単著_日本語	なし	513	88.8	−0.15069	0.011214
	あり	65	11.2	1.189305	−0.08851
発表_人工知能学会	なし	483	83.6	−0.24208	0.092083
	あり	95	16.4	1.230777	−0.46817
発表_情報処理学会	0	443	76.6	−0.30253	0.104184
	Low	70	12.1	0.778326	−0.15139
	High	65	11.2	1.223636	−0.54702
発表_日本機械学会	なし	558	96.5	−0.03649	0.0423
	あり	20	3.5	1.018019	−1.18016
発表_日本知能情報ファジィ学会	なし	566	97.9	−0.01915	0.023704
	あり	12	2.1	0.90314	−1.11806
発表_日本音響学会	なし	542	93.8	−0.07547	0.013077
	あり	36	6.2	1.136285	−0.19688
発表_言語処理学会	なし	492	85.1	−0.17585	0.000115
	あり	86	14.9	1.00602	−0.00066
発表_計測自動制御学会	なし	539	93.3	−0.06776	0.084828
	あり	39	6.7	0.936481	−1.17236
発表_電子情報通信学会	0	458	79.2	−0.27739	0.12943
	Low	63	10.9	1.034221	−0.09284
	High	57	9.9	1.085805	−0.93737
発表_電気関係学会	なし	559	96.7	−0.03603	0.021813
	あり	19	3.3	1.060113	−0.64177
発表_その他	0	391	67.6	−0.37492	0.251443
	Low	103	17.8	0.577643	−0.26465
	High	84	14.5	1.036846	−0.8459
研究費_企業	なし	561	97.1	−0.03548	0.012891
	あり	17	2.9	1.170817	−0.42539
研究費_財団	なし	534	92.4	−0.09047	0.035612
	あり	44	7.6	1.097927	−0.43221
研究費_科研費＜500	0	303	52.4	−0.61038	0.088446
	Low	166	28.7	0.48146	−0.06241
	High	109	18.9	0.96351	−0.15082
研究費_科研費500-2000	0	369	63.8	−0.45699	−0.04038
	Low	129	22.3	0.481591	0.040084
	High	80	13.8	1.331303	0.121597
研究費_科研費＞2000	0	451	78	−0.33507	−0.05322
	Low	96	16.6	0.990708	0.10613
	High	31	5.4	1.806674	0.445569
研究費_その他の共同・競争的資金研究	0	407	70.4	−0.31381	0.055364
	Low	88	15.2	0.59225	0.031993
	High	83	14.4	0.910888	−0.3054
研究費_その他の科研費	0	388	67.1	−0.40682	−0.04687
	Low	106	18.3	0.513033	0.00016
	High	84	14.5	1.23172	0.216271
資歴_国際学会	0	475	82.2	−0.22969	0.067395
	Low	65	11.2	0.833402	−0.24079
	High	38	6.6	1.445561	−0.43056
資歴_大学関連	なし	537	92.9	−0.07315	0.024311
	あり	41	7.1	0.95808	−0.31841
資歴_学会関連	0	325	56.2	−0.51902	0.122697
	Low	133	23	0.334238	−0.27297
	High	120	20.8	1.035231	−0.02976
資歴_その他	0	380	65.7	−0.43067	0.02289
	Low	118	20.4	0.509839	0.063826
	High	80	13.8	1.293688	−0.20287
雑誌_ACL	0	398	68.9	−0.00499	−0.41303
	Low	108	18.7	−0.28299	0.245048
	High	72	12.5	0.452084	1.915581
雑誌_Acoustical Science and Technology	なし	557	96.4	−0.04355	0.003906
	あり	21	3.6	1.155181	−0.10361
雑誌_COLING	0	414	71.6	0.02818	−0.35742
	Low	82	14.2	−0.41746	0.120418
	High	82	14.2	0.275184	1.684101
雑誌_CoRR	なし	484	83.7	−0.17255	−0.09403
	あり	94	16.3	0.888458	0.484173
雑誌_EMNLP	0	467	80.8	−0.0388	−0.27757
	Low	67	11.6	0.113042	0.374073
	High	44	7.6	0.239701	2.376411
雑誌_ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing − Proceedings	なし	554	95.8	−0.04989	−0.00868
	あり	24	4.2	1.151623	0.200472
雑誌_IEEE Access	なし	525	90.8	−0.12302	−0.02125
	あり	53	9.2	1.21863	0.210453
雑誌_IEICE Transactions on Information and Systems	なし	515	89.1	−0.15449	−0.0188
	あり	63	10.9	1.262886	0.153671
雑誌_IEICE TRANSACTIONS ON INFORMATION AND SYSTEMS	なし	484	83.7	−0.22093	−0.07254
	あり	94	16.3	1.137576	0.373498
雑誌_Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)	0	447	77.3	−0.2795	−0.00327
	Low	66	11.4	0.600279	0.229095
	High	65	11.2	1.3126	−0.21016
雑誌_LREC	0	394	68.2	−0.08769	−0.33167
	Low	112	19.4	0.04137	0.19446
	High	72	12.5	0.415497	1.512457
雑誌_LREC-COLING	なし	522	90.3	−0.03996	−0.16225
	あり	56	9.7	0.372476	1.512398
雑誌_NAACL	なし	523	90.5	−0.03675	−0.19803
	あり	55	9.5	0.349438	1.883098
雑誌_Systems and Computers in Japan	なし	514	88.9	−0.17199	0.005755
	あり	64	11.1	1.381306	−0.04622
雑誌_Transactions of the Japanese Society for Artificial Intelligence	なし	528	91.3	−0.12398	−0.04132
	あり	50	8.7	1.309269	0.436332
雑誌_人工知能学会全国大会論文集	なし	520	90	−0.11628	−0.04945
	あり	58	10	1.042523	0.443312
雑誌_人工知能学会誌	なし	514	88.9	−0.16385	−0.02949
	あり	64	11.1	1.315933	0.23681
雑誌_人工知能学会論文誌	0	460	79.6	−0.27267	−0.03301
	Low	71	12.3	0.918594	0.006901
	High	47	8.1	1.281043	0.284149
雑誌_情報処理	なし	536	92.7	−0.10415	−0.0479
	あり	42	7.3	1.329106	0.611282
雑誌_情報処理学会研究報告	なし	558	96.5	−0.06231	−0.03036
	あり	20	3.5	1.73833	0.846962
雑誌_情報処理学会研究報告_自然言語処理研究会報告	なし	565	97.8	−0.03492	−0.02684
	あり	13	2.2	1.517751	1.166691
雑誌_情報処理学会論文誌	0	408	70.6	−0.39643	−0.08896
	Low	85	14.7	0.733333	0.000711
	High	85	14.7	1.169553	0.426297
雑誌_日本ロボット学会誌	なし	556	96.2	−0.0451	−0.01247
	あり	22	3.8	1.13971	0.315134
雑誌_自然言語処理	0	469	81.1	−0.20163	−0.19218
	Low	66	11.4	0.729754	0.231909
	High	43	7.4	1.079079	1.740161
雑誌_電子情報通信学会技術研究報告	なし	549	95	−0.07917	−0.02661
	あり	29	5	1.498861	0.503784
雑誌_電子情報通信学会論文誌	なし	514	88.9	−0.15346	−0.00754
	あり	64	11.1	1.232515	0.060555
雑誌_その他国際学会	0	240	41.5	0.087829	−0.52261
	Low	199	34.4	−0.30285	−0.2258
	High	139	24	0.281934	1.225606
雑誌_その他	0	193	33.4	−0.93563	0.213322
	Low	193	33.4	−0.07583	−0.20287
	High	192	33.2	1.016727	−0.0105
サプリメンタリ変数					
分野_AI人狼	0	441	76.3	−0.09086	0.228101
	1	3	0.5	0.573111	0.022774
	NA	134	23.2	0.286185	−0.7512
分野_LLM	0	440	76.1	−0.08927	0.230572
	1	4	0.7	0.232544	−0.19778
	NA	134	23.2	0.286185	−0.7512
分野_SNS分析	0	438	75.8	−0.08207	0.231283
	1	6	1	−0.4006	−0.10685
	NA	134	23.2	0.286185	−0.7512
分野_トピックモデル	0	441	76.3	−0.09263	0.226454
	1	3	0.5	0.833671	0.264856
	NA	134	23.2	0.286185	−0.7512
分野_フィードバック生成	0	437	75.6	−0.0917	0.23314
	1	7	1.2	0.246381	−0.17447
	NA	134	23.2	0.286185	−0.7512
分野_フランス語（その他言語）	0	443	76.6	−0.09617	0.224733
	1	1	0.2	4.254904	1.104081
	NA	134	23.2	0.286185	−0.7512
分野_マルチモーダル（画像、動画）	0	429	74.2	−0.08271	0.228887
	1	15	2.6	−0.19102	0.164558
	NA	134	23.2	0.286185	−0.7512
分野_固有表現認識、関係認識	0	405	70.1	−0.09837	0.243371
	1	39	6.7	0.038236	0.053731
	NA	134	23.2	0.286185	−0.7512
分野_埋め込みモデル	0	154	26.6	0.09663	0.31677
	1	290	50.2	−0.18355	0.17889
	NA	134	23.2	0.286185	−0.7512
分野_対話	0	414	71.6	−0.1197	0.207635
	1	30	5.2	0.373546	0.489999
	NA	134	23.2	0.286185	−0.7512
分野_広告	0	442	76.5	−0.08549	0.222264
	1	2	0.3	−0.28089	1.209907
	NA	134	23.2	0.286185	−0.7512
分野_文法修正	0	440	76.1	−0.08902	0.232837
	1	4	0.7	0.205158	−0.44684
	NA	134	23.2	0.286185	−0.7512
分野_料理レシピ	0	441	76.3	−0.08499	0.227419
	1	3	0.5	−0.28942	0.122978
	NA	134	23.2	0.286185	−0.7512
分野_日本語処理	0	383	66.3	−0.07938	0.184928
	1	61	10.6	−0.13024	0.489073
	NA	134	23.2	0.286185	−0.7512
分野_構文解析	0	427	73.9	−0.09041	0.222681
	1	17	2.9	0.015119	0.327992
	NA	134	23.2	0.286185	−0.7512
分野_機械翻訳	0	144	24.9	0.053873	0.093224
	1	300	51.9	−0.15369	0.290788
	NA	134	23.2	0.286185	−0.7512
分野_機械翻訳データセット（辞書）	0	423	73.2	−0.11572	0.22881
	1	21	3.6	0.504775	0.184471
	NA	134	23.2	0.286185	−0.7512
分野_物語生成	0	440	76.1	−0.07934	0.230085
	1	4	0.7	−0.8601	−0.14419
	NA	134	23.2	0.286185	−0.7512
分野_自動運転	0	443	76.6	−0.08698	0.229089
	1	1	0.2	0.183418	−0.82559
	NA	134	23.2	0.286185	−0.7512
分野_要約生成	0	422	73	−0.09977	0.223989
	1	22	3.8	0.170583	0.278963
	NA	134	23.2	0.286185	−0.7512
分野_議事録処理	0	441	76.3	−0.08343	0.229522
	1	3	0.5	−0.51868	−0.18623
	NA	134	23.2	0.286185	−0.7512
分野_質問応答	0	431	74.6	−0.09279	0.228187
	1	13	2.2	0.126348	0.177861
	NA	134	23.2	0.286185	−0.7512
分野_金融	0	439	76	−0.08571	0.232972
	1	5	0.9	−0.14418	−0.32278
	NA	134	23.2	0.286185	−0.7512
分野_音声感情分析	0	431	74.6	−0.0935	0.240502
	1	13	2.2	0.150074	−0.23042
	NA	134	23.2	0.286185	−0.7512
分野_音声認識	0	423	73.2	−0.09057	0.234503
	1	21	3.6	−0.00171	0.069807
	NA	134	23.2	0.286185	−0.7512