DEIM2022 C34-1(day3 p17)

既存データセットへの尤度分布による意見文の性質表現

擬似コーパスとしての既存データセット集合の利活用の可能性

竹元　亨舟†　　山西　良典†　　西原　陽子††　　吉田　光男†††　　大須賀智子††††
大山　敬三††††

† 関西大学総合情報学部総合情報学科　〒 569–1095 大阪府高槻市霊仙寺町 2–1–1
†† 立命館大学情報理工学部　〒 525–8577 滋賀県草津市野路東 1–1–1
††† 筑波大学ビシネスサイエンス系　〒 112–0012 東京都文京区大塚 3–29–1
†††† 国立情報学研究所　〒 101–8430 東京都千代田区一ツ橋 2–1–2
E-mail: †{k228105,ryama}@kansai-u.ac.jp, ††nisihara@fc.ritsumei.ac.jp,
†††mitsuo@gssm.otsuka.tsukuba.ac.jp, ††††{osuga,oyama}@nii.ac.jp

あらまし　本稿では，既存データセット集合を擬似コーパスとして利活用することで，分析対象とするテキストデータへのアノテーションを必要としない意見文の性質表現方法を提案する．ソーシャルメディアの登場により，一般ユーザが自由に意見をウェブ上に発信し，他者とのコミュニケーションをとることが容易になった．ウェブ上に存在する様々な意見を分析するための手段としてテキスト分類の応用が考えられる．従来の機械学習手法を用いたテキスト分類では，分類モデルの学習に分析対象となるテキストに対してあらかじめアノテーションが付与された大量のデータを必要とすることが一般的である．しかしながら，時代や社会通念，個人の感性によって分類ラベルの基準は異なる可能性が考えられるため，それらの違いに応じたアノテーションを用意しなければ有用性の高いテキスト分類モデルの構築はできない．このためのアノテーション付きコーパスの準備には人的・時間的コストを要する．提案手法では，性質が異なる文書をそれぞれ含む既存データセットを複数組み合わせて擬似コーパスとする．この擬似コーパスを学習し，各データセットへの分類モデルを構築する．既存データセットへの尤度分布によってテキストの性質を表現し，アノテーション作業を必要としないテキスト分類の可能性を検討した．

キーワード　感情分析，SNS 分析，テキスト分類，既存データセットの利活用，擬似コーパス，意見文の性質表現

1　はじめに

ソーシャルメディアの登場により，誰もがウェブ上に意見を発信できるようになった．現在のウェブはさまざまな情報で溢れかえっており，ユーザは自身が求める情報を適切に手に入れるために情報を整理・分類する必要がある．ウェブ上の情報の増加に伴ってソーシャルメディア上のテキスト分類技術へのニーズは高まりを見せている．

機械学習アプローチによるテキスト分類では，テキストと対応するラベルの対から構成されるコーパスを用意して学習することで，用意されたラベルについての分類モデルを学習する．このとき，従来の研究ではあらかじめ用意されたラベルは単一の概念を示す離散的なシンボルのラベルである場合が多い．例えば，不適切な表現か否かを分類する場合には，あるテキスト集合に対して「不適切である」か「不適切ではない」かの 2 種類のラベルをアノテートしたコーパスが用いられる．しかし，それらのアノテートされたラベルは，個々人の趣味嗜好によらず普遍的であり，時代や社会の変化に対しても不変であるとは限らない．

アノテーションラベルの普遍性を担保するためには，複数のラベラーからのアノテーションを統計的に扱うことで対応できる可能性がある．一方で，上記の例で言えば，どのような文を「不適切な表現」とするかの基準は時代や社会の風潮などにより変化する可能性がある．時代や社会に応じて変化する基準や感性は固定的に与えられたアノテーションラベルを用いることでは対応できず，変化に応じた基準で用意されたアノテーションが付与されたコーパスが新たに必要となる．これは，アノテーション作業を必要とするコーパスを利用した機械学習アプローチが潜在的にもつテキスト分類の共通課題であると考える．

1.1　関連研究

ソーシャルメディアにおけるテキスト情報の分類に関連する研究や実施例については様々なアプローチが報告されている．感情分析の観点からは，投稿文に含有される特徴量の抽出 [1] や感情の正負に基づいた分類 [2] が行われている．近年の代表的研究としては，Wang らは感情分析モデルに対して Attention 機構を取り入れることで，推定性能の向上を図っている [3]．学習機構自体の改善については多くの報告があるものの，学習データの準備やアノテーションラベルの表現方法に関する研究報告は少ない．

ソーシャルメディア上の膨大な情報に様々な不適切情報が

含まれることも問題となっており，この問題に対しても様々なアプローチ（例えば，文献 [4]）が提案されている．最近では，ソーシャルメディアのプラットフォーム側が，不適切な表現や攻撃的な表現について閲覧せずに済む方法 [1]や投稿時に再検討を促す仕組み [2]なども提供している．これらのアプローチでは，どのような情報が不適切であるのかをアノテートしたコーパスを学習データとして準備する必要がある．このための学習データの準備に関しては，Kim ら [5] は，Kaggle 社の提供する質問サイト上の投稿の誠実度データセット [3]を整理して，不適切投稿の学習データとして利用している．

本稿では，普遍的なデータセットの集合を疑似コーパスとして利用することで，専用の学習コーパスを準備しなければならない問題の解決をねらう．また，擬似コーパス中の各データセットへの尤度分布をテキストの性質表現とすることで，社会通念や個人に応じて異なるアノテーションラベルに対する感度や感性の違いに柔軟な学習モデルの構築を目指す．

1.2 本稿の貢献

提案手法では，学習用のコーパスとしてあらかじめアノテーションを付与したデータを用意するのではなく，複数の既存データセットから擬似コーパスを構成することで，アノテーション作業を必要としないテキスト情報の性質表現を提案する．提案手法によって構築した分類モデルは，与えられた文に対応するアノテーションラベルではなく，各データセットへの尤度分布を推定する．この尤度分布を入力テキストに対する擬似的なソフトラベルとして扱い，その分布を解釈することでテキスト情報の性質表現を試みる．

擬似コーパスを構築するためには，分類対象であるテキストの性質を表現すると期待される異なる性質をもつ複数のデータセットが必要となる．国立情報学研究所の情報学研究データリポジトリ（IDR）[6] には多種多様なデータセットやコーパスが集約されている．これらのデータセットは特定のドメインでの課題解決に用いられることがほとんどであったが，提案手法では擬似コーパス構築に利活用する．3 種類の既存データセットを混合した擬似コーパスを構築し，ソーシャルメディアに投稿された意見を分類するモデルを構築する．このモデルにより得られた擬似コーパスへの尤度分布を,「立場」「論理性」「文体」という観点から解釈し，提案手法による文書の性質表現の有用性を議論する．

2 提 案 手 法

提案手法では，分析対象となるドメインの学習用データに対してアノテーション作業を行わずにテキスト分類モデルを構築する．近年の機械学習アプローチによるテキスト分類では，単語分散表現 [7] が用いられることが多い．単語分散表現は，ある単語と周辺語との関係性から単語を高次元ベクトル上に埋め込むことで獲得される特徴量の表現手法である．単語分散表現を用いたテキスト分類モデルでは，ラベルごとにこれらの特徴量がどのような分布であるかを学習する．提案手法では，目的ドメインのテキストの各性質を表現する特徴量の分布と類似した分布を持つと考えられるドメインのデータセットを用意する．用意したデータセットのドメインを分類するモデルを構築し，目的ドメインの入力テキストに対してドメイン推定を行うことで，結果的に既存の各データセットの性質に対する尤もらしさをしめす分布として，入力テキストの性質表現が得られると考える．

山西らによる提案 [8] では，分類目的のラベルそれぞれに対して，直接的に対応するデータセットを用意している．レストランのレビューの観点分類を課題として取り上げ，調理レシピデータセットを料理の観点に，宿泊施設のレビューデータセットをホスピタリティの観点に，それぞれ対応させて学習し，アノテーション作業を必要としないテキスト分類の基本的アイデアを示した．本稿では，扱うテキストの「ラベル」に関する捉え方をさらに「性質」へと分解した．既存データセットを学習することで得られた分類モデルによって推定された入力テキストの各既存データセットへの尤度分布はすなわち，各データセットらしさを示すことになる．この尤度分布自体は各データセットらしさを示しているだけであるため，時代や世情の変化によらず不変的な推定モデルとなる．得られる分布形状を時代や世情に応じて解釈することで，分析対象のドメインに対する人間の認知や理解の変化に対応可能なテキスト分類が可能になると考えられる．

2.1 提案手法の処理手順

提案手法の処理は，以下のように整理できる．

（1） 学習する各データセットのテキストを正規化，形態素解析して整形

（2） 学習する各データセットのテキストを混合し，擬似コーパスを構築

（3） 擬似コーパスを train，test，valid データへと分割

（4） fastText モデルのファインチューニング

データセット中のテキストには機械学習の妨げになるノイズが含まれるため，それらを正規化処理によって取り除く．正規化処理では主に以下の処理を行なう．

- URL やメールアドレスなど，参照の除去
- 顔文字や絵文字のような装飾文字の除去
- 英数字の全角文字，繰り返し表現，数値の正規化
- 句読点を除く記号，環境依存文字の除去

日本語形態素解析システムである MeCab [9] で NEologd 辞書を用いて形態素解析を行う．各データセットからそれぞれ同一件数のデータをランダムに抽出し，それぞれ対応するデータセットを示すラベルを付与することで擬似コーパスを構築する．用意した擬似コーパスを用いてモデルの学習，およびファインチューニングを行って分類モデルを構築する．分類モデルには，

1：https://about.instagram.com/blog/announcements/introducing-new-tools-to-protect-our-community-from-abuse（2021/12/17 確認）

2：https://blog.twitter.com/en_us/topics/product/2021/tweeting-with-consideration（2021/12/24 確認）

3：https://www.kaggle.com/c/quora-insincere-questions-classification/overview（2021/12/29 確認）

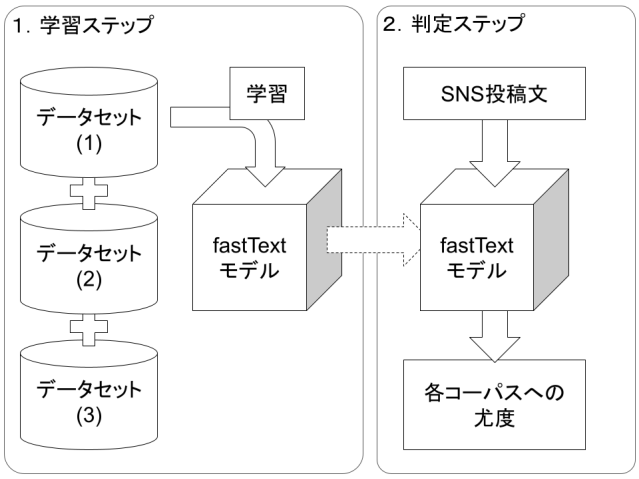

図 1: 3 ラベル分類手法の概要

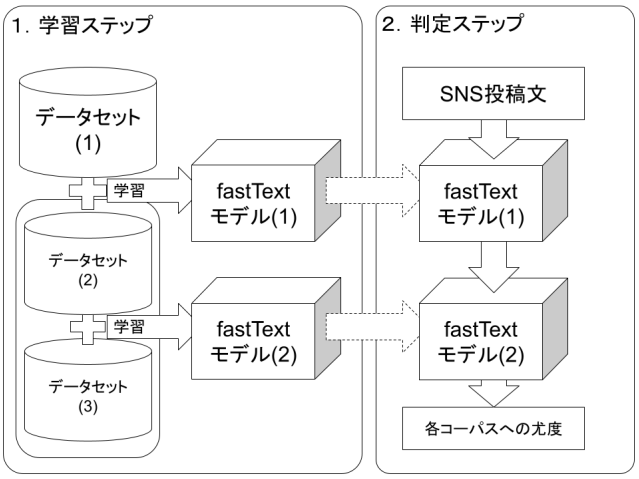

図 2: 2 段階分類手法の概要

分類ドメインに適応的に文書分散表現を獲得する fastText [10] を用いた．fastText は Skip-gram 法，もしくは CBOW 法を用いて文中の単語を高次元ベクトル化し，意味空間への埋め込みにより分散表現を獲得する．今回は，より高い精度が示されている Skip-gram 法を用いた．ただし，近年の研究では，CBOW の性能を引き出す研究が実施されている [11] ため，分類モデルの検討は今後の課題とする．

2.2 3 種類のデータセットを用いた提案手法の適用方法

本稿ではソーシャルメディア上への投稿文の性質を分類するために，3 種類のデータセットを組み合わせた擬似コーパスを用いる．ソーシャルメディアへの投稿文に対して，これらのデータセットへの尤度分布を推定するうえでは，以下の 2 種類のアプローチが考えられる．

（1） 3 ラベル分類手法

（2） 2 段階分類手法

どちらのアプローチでも，あらかじめ 3 種類の既存データセットを組合わせて構築した擬似コーパスによって分類モデルを構築する．この分類モデルに対して，ソーシャルメディアの投稿文を入力とすることで，入力された投稿文を表現する fastText の分散表現が得られる．この分散表現に従って，擬似コーパスを構成する各データセットに対しての尤もらしさを示す分布が出力として得られる．3 ラベル分類手法と 2 段階分類手法では，分類モデルの数と推論過程が異なる．各手法の詳細は，2.2.1 節と 2.2.2 節でそれぞれ示す．

2.2.1 3 ラベル分類手法

図 1 に，3 ラベル分類手法の概要を示す．本手法は学習ステップと判定ステップの 2 段階で構成される．

学習ステップでは，各データセットをそれぞれ学習用のラベルとして扱い，一度に 3 種類のデータセットドメインの分類モデルを構築する．3 種類のデータセットのそれぞれの性質の違いが 1 つの分類モデルで学習されることになる．

判定ステップでは，学習したモデルに分析対象とするドメインの入力テキストを判定させる．入力テキストのノイズによるモデルの精度への影響を減らすため，分析対象となる入力テキストにも 2.1 節で行なった正規化処理と同様の処理を施す．判定の結果，擬似コーパスを構成する各データセットに対する入力テキストの各尤度が分布として得られる．このアプローチでは分類因子として，3 種類のデータセット間でのテキストの特徴分布の差が学習されることが期待される．

2.2.2 2 段階分類手法

図 2 に，2 段階分類手法の概要を示す．本手法においても，前述の 3 ラベル分類手法と同様に，大きくは学習ステップと判定ステップに分けられる．3 ラベル分類手法とは，学習ステップにおける分類モデルの構成と分類モデルの数が異なる．

学習ステップでは，2 種類のデータセットの組み合わせ間で大きく異なると考えられる性質に着目し，それぞれの性質に対して段階的に分類するモデルを構築する．1 段階目では，3 種類のデータセットを性質的に 1:2 に分割し，1 つのデータセットと 2 種類のデータセットの集合との分類モデルを構築する．2 段階目では，1 段階目でまとめた 2 種類のデータセットを分類するモデルを構築する．この 2 段階の二項分類を経ることで，各モデルが分類するべき因子がより明確になり，優れた結果が期待される．

判定ステップにおいても，2 段階のフェーズに分割して分類が行われる．各フェーズでは，それぞれ 2 種類のデータセットを学習した分類モデルが構築される．第 1 フェーズでの分類の結果，2 種類のデータセットの集合への尤度が高かったものについては，第 2 フェーズでの分類モデルでの分類に進むことで，最終的な分類結果を得る．本アプローチでは，各データセットに期待するテキストの性質についての仮説に従って，分類基準を意図的に限定しながら，人間が文書を読み解く認知段階に即した分類を行っていくことになる．

3 実験設定

実験では，3 種類のデータセットが組み合わされて構成された擬似コーパスを学習した分類モデルを構築した．この分類モデルから推定された各データセットへの尤度分布によって，ソーシャルメディア上の意見の性質表現を行った．

表 1: L，F，Cの各データセットに期待される性質

データセット	内容	立場	論理性	文体
L	弁護士相談	中立的	論理的	文語的
F	不満投稿	批判的	感情的	文語的
C	会話	好意的	感情的	口語的

3.1 分析対象のデータ

分類対象とするソーシャルメディアのデータは Twitter より，緊急事態宣言が発令された 2021/08/11 の 10:00-23:00 に日本人政治家のツイートに対して送信された日本語のリプライ投稿のテキスト（以下，リプライ）10,663 件を使用した．リプライの多くは，緊急事態宣言を発令した政府に対する主張や不満などが含有される意見である．

これらの投稿を提案手法による分類タスクによる尤度分布から解釈し，主張の性質を考察する．分析対象としたリプライ投稿に含まれる意見の性質は，不適切であるか否かの単純な二項問題ではない．例えば，一見すると過激な単語を用いていても，主張に論理的な一貫性が見られれば，議論の可能性がある．実験では，主張の攻撃性や内容の密度，意見の伝わりやすさの観点から，それぞれのリプライに含まれる単語や表現 [12] を考察する．

3.2 学習に用いる既存データセット

3.1 節で設定した観点での分類と考察を行うために，擬似コーパスを構成する既存データセットとして,「立場」「論理性」「文体」の観点で異なる単語や表現が用いられると考えられる 3 種類のデータセットを検討し，NII が提供する弁護士ドットコムデータ [13] と不満調査データ [14]，千葉大学 3 人会話コーパス [15] [16] の書き起こし文を採用した．

これらのデータセットの性質を，分析対象データの性質に照らし合わせ,「立場」「論理性」「文体」の観点から表 1 のように仮定する．弁護士ドットコムデータ（以下データセットL）は弁護士ドットコム株式会社の運営するオンライン法律相談サービス「みんなの法律相談」に投稿された約 25 万件の質問と，それに対する弁護士の回答テキストである．不満調査データ（以下データセットF）は株式会社 Insight Tech が運営する Web サービス「不満買取センター」に投稿された約 525 万件の不満のテキストである．千葉大学 3 人会話コーパスの書き起こし文（以下データセットC）は同性 3 人からなる友人同士 12 組が行なった雑談，計約 2 時間分を文字に起こしたものである．

3.3 提案手法の適用

擬似コーパスを構築するため，L，F，Cの各データセットからそれぞれ 2,500 件のテキストをランダムに抽出し，それぞれ対応するデータセットを示すL，F，Cのラベルを付与する．データセットCに含まれる文が 2,500 件程度と他コーパスに比べて少なく，各ラベルのデータ件数の不均衡を回避するため，ダウンサンプリングした．擬似コーパスは 6 割を train，2 割を test，残った 2 割を validation データと設定し，無作為に分割する．これらの擬似コーパスを用いてモデルの学習，およびファインチューニングを行い，その精度を評価する．

モデルのファインチューニングについては，すべての品詞を対象とした学習モデルでのメタパラメータを用いることとした．3 ラベル分類モデルでは，ファインチューニングを行った結果，$epoch = 25$，$lr = 0.5$，$wordNgrams = 2$ の時に，適合率 0.976，再現率 0.976 と高い精度を示した．このパラメータで学習したモデルに対し test データを予測させたところ，適合率と再現率共に 0.978 の精度を確認し，このパラメータが適切であると判断した．

データセットL，Fの 2 種類とデータセットCの間では,「文体」という性質が大きく異なると考察される．そこで，2 段階分類手法では，第 1 フェーズでは文の「文体」に，第 2 フェーズでは文の「立場」「論理性」の違いに着目した分類を期待する．以下の 2 つの fastText 分類モデルを学習し，分類対象のテキストを各モデルを順に用いて分類を行う．

（1） L，Fの混合ラベルとCラベルを用いた分類モデル

（2） LラベルとFラベルを用いた分類モデル

各フェーズの分類モデル学習では，validation データへの分類精度をもとに，それぞれ異なるハイパーパラメータでファインチューニングを行った．第 1 フェーズでは $epoch = 25$, $lr = 0.9$, $wordNgrams = 1$ で行い，適合率 0.990，再現率 0.990 を，第 2 フェーズでは $epoch = 50$, $lr = 0.5$, $wordNgrams = 2$ で行い，適合率 0.981，再現率 0.981 を示した．これらを各フェーズでの分類モデルのハイパーパラメータとして設定した．

4 実験結果

表 2 に，3 ラベル分類モデルと 2 段階分類モデルでの性質表現の結果の一部を示す．ただし，表には分析対象のリプライの本文から自立語のみを取り出して示す．4.1 節と 4.2 節に，それぞれ 3 ラベル分類モデルと 2 段階分類モデルでの結果の考察を示す．

4.1 3 ラベル分類モデルについての考察

分類結果におけるL，F，Cの各ラベルの尤度の平均は 0.307, 0.429，0.264 となった．分散は 0.169，0.190，0.169 となり，どのラベルにおいても大きな広がりは見られない．結果の尤度分布は多くがバスタブ型となっており，Fラベルの尤度が 0 となるものが 440 件存在し，L（595 件），C（681 件）よりも少ない．これは，分析対象が緊急事態宣言が発令された日に政治家の投稿に対してのリプライであり，現状への不満を示すものが多かったためであると考えられる．この結果から，分析対象の投稿テキストは全体的にFラベルに属する傾向が強く，Cラベルの傾向が弱いことが伺える．

各ラベルに属するテキストの特徴を分析するため，各ラベルの尤度の高いテキスト中に現れる単語の特徴を調べた．Lラベルの尤度 1.000 を示したテキストには丁寧な表現，感謝を表す表現（ありがとう，よろしくお願いなど）が多くみられた．文末（ます，でしょうか）や単語（ご挨拶，お伝え）についても丁寧な表現が用いられ，カタカナの表現も多くみられた（ステ

表 2: 3 ラベル分類モデルと 2 段階分類モデルを用いて分類を行なった結果の例．表には本文中の自立語のみを示し，表中の [数] は数字を正規化した表記を示す．分類結果の各ラベルの尤度は小数第四位で四捨五入したものである．

		3 labels			2 phases			
					1/ 2 phases		2/ 2 phases	
	本文中の自立語	L	F	C	L F	C	L	F
0	宇宙人	0.000	0.000	1.000	0.000	1.000	0.000	1.000
1	さん, 原口, 大丈夫	1.000	0.000	0.000	1.000	0.000	1.000	0.000
2	さ, すぎる, 変わる, 対応, 無い, 節操	0.000	1.000	0.000	1.000	0.000	0.000	1.000
3	[数], [数], いる, する, ない, ふざける, やる気, ワクチン, ワクチン, 丁寧, 予約, 人, 人, 休業, 会社, 取れる, 口, 増える, 多く, 守る, 安倍前総理, 悪意, 打つ, 打つ手, 政府, 早い, 早い, 焦る, 焦る, 番, 破れ, 程度, 菅総理, 馬鹿者	0.001	0.997	0.002	1.000	0.000	0.000	1.000
4	これだけ, ダメ, ロックダウン, 抑える, 立憲	0.000	0.367	0.633	0.860	0.140	0.000	1.000
5	する, 人, 何とか, 家, 恥ずかしい, 支持率, 自分, 言う	0.000	0.087	0.914	0.029	0.971	0.000	1.000
6	ある, する, ない, コロナ, ステイホーム, 大臣, 恵まれる, 笑, 罹る, 言う, 身内	0.999	0.000	0.000	1.000	0.000	1.000	0.000
7	[数], する, 対応, 早く, 遅い, 類	0.000	1.000	0.000	1.000	0.000	0.000	1.000
8	ある, する, そう, 中等, 可能性, 感染, 感染, 時間, 水疱瘡, 水痘, 短い, 部屋, 飛行機	0.063	0.937	0.000	1.000	0.000	0.162	0.838
9	信じる, 捏造	0.324	0.000	0.676	0.001	0.999	1.000	0.000
10	CEO, スーザンウォジスキ, 氏	0.000	0.000	1.000	0.000	1.000	0.000	1.000
11	[数], BlaZe, すごい, つく, ショップ, スケボー, ボード, 三連休, 亀戸, 代, 効果, 在庫, 多く, 子, 底, 来る, 来店, 波, 者, 若い, 達, 金メダル, 陳列	0.005	0.995	0.000	1.000	0.000	0.001	0.999
12	する, 真似	0.000	0.019	0.981	0.000	1.000	0.000	1.000
13	[数], [数], する, ない, ない, ひるおび!, 五輪, 五輪, 五輪, 北村, 原因, 反す, 専門家, 専門家, 強行, 感染症, 政府, 本日, 氏, 民放, 波, 波, 結果, 義浩, 重要, 開催, 開催	0.878	0.122	0.000	1.000	0.000	0.918	0.082
14	PCR, する, キャンセル, 人, 受ける, 受ける, 受け手, 大丈夫, 強制, 思う, 旅行, 行く, 陽性, 頭	0.091	0.597	0.311	0.930	0.070	0.019	0.981

イホーム，パラリンピックなど）．Fラベルの尤度 1.000 を示したテキストには具体的な主張に関連する内容を示す表現（給与，国民，勘弁など）が多くみられた．同時に，乱暴な表現も見られた（若作り，ヘド，うるせーよ，ブスなど）．Cラベルの尤度 1.000 を示したテキストには短い表現，内容のない表現（あれですね，意味不明，ないわー，ごくろうさん），呼びかけ（あんた，先生，あなた）などが見られた．これらの特徴から，各ラベル間の分類因子として，投稿テキストの「丁寧さ」と「会話らしさ」が大きく影響しており，丁寧な表現が含まれるかどうかによってLラベルかそうでないかに分類され，内容の密度によってCラベルかそうでないかに分類されると考えられる．また，Fラベルと推定されたリプライには主張を断言し，相手の返事を期待しないテキストが多く，Cラベルでは能動的な情報発信が見られず，受動的な受け答えが多いように見受けられた．

複数のラベルが組み合わされた際の性質を考察する．3 ラベル中 2 ラベルの値が近く，そのかけ合わせた値が 0.230 以上で，残りの数値がほぼ 0 に近いものをピックアップして分析対象として抽出した．L，Fラベルの組み合わせは 30 件，FとCでは 16 件，LとCでは 3 件の投稿が検出された．L，Fラベルの組み合わせは丁寧な表現で不満を表しており，これらは意見として相手に受理される可能性が高いと考えられる．F，Cラベルの組み合わせは内容のない言葉で不満を表しており，これらは相手にとって有意な意見として受け取ってもらえる可能性は低い．一方で，L，Cラベルの組み合わせはほとんどなかった．これはLとCの両方のコーパスで使われる単語集合に共通部分がないためと考えられる．全てのラベルが同程度の尤度を持つ投稿は 2 件しか確認されず，どのラベルにも分類ができないものは少なかった．

4.2 2 段階分類モデルについての考察

2 段階分類モデルにおいても 3 ラベル分類と同様に，バスタブ型の尤度分布が多くみられた．各分類において，あるラベルへの尤度が 0.500 より大きいテキストをそのラベルに属するとして考察した．

第 1 フェーズでは，LとFの混合ラベル（以下LFラベル）とCラベルの二項分類を行なった．LFラベル，Cラベルに属するテキストは，それぞれ形態素数の平均が 32.817 と 9.853 であり，大きな差が見られた．LFラベルは，Cラベルに比べて長い文章となっており，より内容の濃い主張であったため，このような差が見られたと考えられる．LFラベルに対して尤度 1.000 を示したテキストは，具体的な提案を含んだ意見や実際の数値を基にした議論などが多かった．Cラベルの尤度 1.000 を示したテキストは，口語調の短文で会話の受け答えとしては成立するものの，意見を読みとることが困難な記述が多く見られた．

表 3: 名詞のみで学習を行なったモデルの分類結果との比較．表 2 と同様に本文は自立語のみを表示し，各ラベルの尤度は小数点第四位で四捨五入している．表中の [数] は数字を正規化した表記を示す．

		3 labels			noun only		
	本文中の自立語	L	F	C	L	F	C
0	宇宙人	0.000	0.000	1.000	0.000	0.004	0.996
1	さん, 原口, 大丈夫	1.000	0.000	0.000	0.006	0.000	0.993
2	さ, 対応, 節操	0.000	1.000	0.000	0.037	0.702	0.260
3	[数], [数], やる気, ワクチン, ワクチン, 丁寧, 予約, 人, 人, 休業, 会社, 口, 多く, 安倍前総理, 悪意, 打つ手, 政府, 番, 破れ, 程度, 菅総理, 馬鹿者	0.001	0.997	0.002	0.093	0.900	0.007
4	これだけ, ダメ, ロックダウン, 立憲	0.000	0.367	0.633	0.051	0.789	0.159
5	人, 何とか, 家, 支持率, 自分	0.000	0.087	0.914	0.002	0.993	0.005
6	コロナ, ステイホーム, 大臣, 笑, 身内	0.999	0.000	0.000	0.048	0.124	0.828
7	[数], 対応, 早く, 類	0.000	1.000	0.000	0.079	0.919	0.002
8	そう, 中等, 可能性, 感染, 感染, 時間, 水疱瘡, 水痘, 部屋, 飛行機	0.063	0.937	0.000	0.213	0.718	0.068
9	捏造	0.324	0.000	0.676	0.013	0.065	0.922
10	CEO, スーザンウォジスキ, 氏	0.000	0.000	1.000	0.000	0.006	0.994
11	[数], BlaZe, ショップ, スケボー, ボード, 三連休, 亀戸, 代, 効果, 在庫, 多く, 子, 底, 来店, 波, 者, 達, 金メダル, 陳列	0.005	0.995	0.000	0.214	0.769	0.017
12	真似	0.000	0.019	0.981	0.000	0.001	0.999
13	[数], [数], ひるおび!, 五輪, 五輪, 五輪, 北村, 原因, 専門家, 専門家, 強行, 感染症, 政府, 本日, 氏, 民放, 波, 波, 結果, 義浩, 重要, 開催, 開催	0.878	0.122	0.000	0.104	0.875	0.022
14	PCR, キャンセル, 人, 受け手, 大丈夫, 強制, 旅行, 陽性, 頭	0.091	0.597	0.311	0.181	0.311	0.509

表 4: 名詞を除外して学習を行なったモデルの分類結果との比較．表 2 と同様に本文は自立語のみを表示し，各ラベルの尤度は小数点第四位で四捨五入している．分析対象となる自立語が存在しなかったものは，本文が空行となっている．

		3 labels			without noun		
	本文中の自立語	L	F	C	L	F	C
0		0.000	0.000	1.000	0.147	0.000	0.853
1		1.000	0.000	0.000	1.000	0.000	0.000
2	すぎる, 変わる, 無い	0.000	1.000	0.000	0.000	1.000	0.000
3	いる, する, ない, ふざける, 取れる, 増える, 守る, 打つ, 早い, 早い, 焦る, 焦る	0.001	0.997	0.002	0.004	0.985	0.011
4	抑える	0.000	0.367	0.633	0.002	0.034	0.964
5	する, 恥ずかしい, 言う	0.000	0.087	0.914	0.003	0.002	0.995
6	ある, する, ない, 恵まれる, 罹る, 言う	0.999	0.000	0.000	0.997	0.000	0.003
7	する, 遅い	0.000	1.000	0.000	0.000	1.000	0.000
8	ある, する, 短い	0.063	0.937	0.000	0.266	0.733	0.001
9	信じる	0.324	0.000	0.676	0.719	0.000	0.281
10		0.000	0.000	1.000	0.029	0.000	0.971
11	すごい, つく, 来る, 若い	0.005	0.995	0.000	0.045	0.955	0.000
12	する	0.000	0.019	0.981	0.593	0.044	0.363
13	する, ない, ない, 反す	0.878	0.122	0.000	0.964	0.036	0.000
14	する, 受ける, 受ける, 思う, 行く	0.091	0.597	0.311	0.118	0.015	0.866

4.3 学習対象とする品詞についての考察

提案手法において構築したモデルが，何を手がかりとして分類を行っているのかについて考察する [17]．例えば，分類モデルが名詞に依存している場合，法律用語が登場する文章では弁護士の投稿の分布であるＬラベルに分類されやすくなる可能性がある．学習対象の品詞による推定結果の違いを分析するために，3 ラベル分類モデルにおいて，学習対象とする品詞を限定した以下の 2 種類のモデルを追加で用意した．

（1） 名詞のみを対象とした学習モデル

（2） 名詞以外を対象とした学習モデル

表 3 に名詞のみを対象とした学習モデル，表 4 に名詞以外を対象とした学習モデルの結果をそれぞれ示す．

これらのモデルの出力はそれぞれ，分類の因子として名詞，文体がどの程度影響を与えているのかを示していると考えられる．結果，先の分類に用いた 10,663 件のツイートのうち，名詞のみを学習した分類モデルでは 5,915 件が，名詞以外を学習した分類モデルでは 8,523 件が先の分類結果と異なるラベルであると判定された．これらの結果から，提案手法による分類モ

デルでは名詞や文体のどちらかに依存して分類されたのではなく，ラベルごとの名詞と文体の組み合わせの特徴が学習されたことが示唆された．

5 おわりに

本稿では，擬似コーパスを用いたテキスト分類によるテキストの性質表現手法を提案した．提案手法では，異なるドメインの既存データセットを組み合わせて擬似コーパスとして用いた．これにより，分析対象である任意のドメインのテキストに対して，アノテーションを必要とせずにテキスト分類ができる可能性を示した．実験では，ソーシャルメディアへ投稿されたテキストの性質に対し,「立場」「論理性」「文体」の観点からテキストを分析するために，弁護士ドットコムデータと不満調査データ，千葉大学 3 人会話コーパスを用いて擬似コーパスを構築した．その結果，各ラベルに割り当てられたテキスト群には，それぞれ特徴的な単語や表現が含まれることが確認された．

今後は，提案手法によって表現された尤度分布をもとにクラスタリングすることでアノテーションに依存しないテキスト分類手法の可能性を検討する．また，他の既存データセットの組み合わせによって，ソーシャルメディア上の意見を本稿とは異なる観点で性質表現できるかについても検討する．

謝　辞

本研究では，国立情報学研究所の IDR データセット提供サービスにより株式会社 Insight Tech から提供を受けた「不満調査データセット」および弁護士ドットコム株式会社から提供を受けた「弁護士ドットコムデータセット」，国立情報学研究所音声資源コンソーシアムから提供を受けた「千葉大学 3 人会話コーパス（Chiba3Party）」をそれぞれ利用した．本研究は，一部，2021 年度国立情報学研究所公募型共同研究（#21S0501），および，JSPS 科研費 JP20K02642 の助成のもと行われた．記して謝意を表す．

文　献

[1] Quanzhi Li, Sameena Shah, Rui Fang, Armineh Nourbakhsh, and Xiaomo Liu. Tweet sentiment analysis by incorporating sentiment-specific word embedding and weighted text features. In *Proceedings of the 2016 IEEE/WIC/ACM International Conference on Web Intelligence*, pp. 568–571, 2016.

[2] Sheeba Naz, Aditi Sharan, and Nidhi Malik. Proceedings of the sentiment classification on twitter data using support vector machine. In *Proceedings of the 2018 IEEE/WIC/ACM International Conference on Web Intelligence*, pp. 676–679, 2018.

[3] Yequan Wang, Minlie Huang, Xiaoyan Zhu, and Li Zhao. Attention-based LSTM for aspect-level sentiment classification. In *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing*, pp. 606–615, 2016.

[4] Cédric Maigrot, Vincent Claveau, and Ewa Kijak. Fusion-based multimodal detection of hoaxes in social networks. In *Proceedings of the 2018 IEEE/WIC/ACM International Conference on Web Intelligence*, pp. 222–229, 2018.

[5] Do Yeon Kim, Xiaohang Li, Sheng Wang, Yunying Zhuo, and Roy Ka-Wei Lee. Topic enhanced word embedding for toxic content detection in Q & A sites. In *Proceedings of the 2019 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*, pp. 1064–1071, 2019.

[6] 国立情報学研究所情報学研究データリポジトリ. `https://www.nii.ac.jp/dsc/idr/index.html`.

[7] Tomas Mikolov, Ilya Sutskever, Kai Chen, Greg S Corrado, and Jeff Dean. Distributed representations of words and phrases and their compositionality. In *Proceedings of the Advances in Neural Information Processing Systems 26*, pp. 3111–3119, 2013.

[8] 山西良典, 藤岡寛子, 西原陽子. 擬似コーパスを用いた飲食店レビューの観点の自動分類. 人工知能学会論文誌, Vol. 36, No. 1, pp. WI2–A_1–8, 2021.

[9] Taku Kudo, Kaoru Yamamoto, and Yuji Matsumoto. Applying conditional random fields to Japanese morphological analysis. In *Proceedings of the 2004 Conference on Empirical Methods in Natural Language Processing*, 2004.

[10] Piotr Bojanowski, Edouard Grave, Armand Joulin, and Tomas Mikolov. Enriching word vectors with subword information. *Transactions of the Association for Computational Linguistics*, Vol. 5, pp. 135–146, 2017.

[11] Ozan İrsoy, Adrian Benton, and Karl Stratos. Corrected CBOW performs as well as skip-gram. In *Proceedings of the Second Workshop on Insights from Negative Results in NLP*, pp. 1–8, 2021.

[12] Carol M. Eastman. Establishing social identity through language use. *Journal of Language and Social Psychology*, Vol. 4, No. 1, pp. 1–20, 1985.

[13] 弁護士ドットコム株式会社. 弁護士ドットコムデータセット. 国立情報学研究所情報学研究データリポジトリ. (データセット), 2020. `https://doi.org/10.32130/idr.12.1`.

[14] 株式会社 InsightTech. 不満調査データ. 国立情報学研究所情報学研究データリポジトリ. (データセット), 2017. https://doi.org/10.32130/idr.7.1.

[15] 伝康晴, 榎本美香. 千葉大学 3 人会話コーパス（Chiba3Party）. 国立情報学研究所音声資源コンソーシアム. (データセット)., 2014. `https://doi.org/10.32130/src.Chiba3Party`.

[16] Yasuharu Den and Mika Enomoto. *A scientific approach to conversational informatics : Description, analysis, and modeling of human conversation*, pp. 307–330. John Wiley & Sons., Hoboken, NJ, 2007.

[17] Mitsuo Yoshida, Takeshi Sakaki, Tetsuro Kobayashi, and Fujio Toriumi. Japanese conservative messages propagate to moderate users better than their liberal counterparts on twitter. *Scientific Reports*, Vol. 11, No. 19224, 2021.