

# 複数の属性の関連性に着目したソーシャルメディアからの市民意見抽出

石田 哲也<sup>†</sup> 関 洋平<sup>††</sup> 柏野和佳子<sup>†††</sup> 神門 典子<sup>††††,†††††</sup>

<sup>†</sup> 筑波大学情報学群知識情報・図書館学類 〒305-8550 茨城県つくば市春日 1 丁目 2 番地

<sup>††</sup> 筑波大学図書館情報メディア系 〒305-8550 茨城県つくば市春日 1 丁目 2 番地

<sup>†††</sup> 国立国語研究所 〒190-8561 東京都立川市緑町 10 丁目 2 番地

<sup>††††</sup> 国立情報学研究所 〒101-8430 東京都千代田区一ツ橋 2 丁目 1 番地 2 号

<sup>†††††</sup> 総合研究大学院大学 〒101-8430 東京都千代田区一ツ橋 2 丁目 1 番地 2 号

E-mail: <sup>†</sup>s1711482@s.tsukuba.ac.jp, <sup>††</sup>yohei@slis.tsukuba.ac.jp, <sup>†††</sup>waka@ninja.ac.jp, <sup>††††</sup>kando@nii.ac.jp

**あらまし** 行政の政策や接客業のサービスの質を向上させるためには、都市で暮らす市民によるフィードバックが重要となる。本研究では、ソーシャルメディアのつぶやきを複数の観点から分析することで、つぶやきに現れる多様な市民意見を整理して抽出する。この際、つぶやきの意見タイプ、地域依存性、極性といった複数の属性を、BERT を用いて推定する。しかし、各属性における選択枝の数や選択枝ごとのデータの割合は異なるため、全ての属性を同様の手法によって高精度で推定するのは困難である。本研究では、各属性の関連性に着目し、BERT を用いた Multi-task learning 手法による各属性の推定精度向上を試みる。さらに、推定した属性を利用して時系列ごとの市民意見の出現頻度を分析することで、市民意見と社会情勢や自治体の対応との相関を確認する。

**キーワード** Twitter, 意見分析, アプレイザル理論, BERT, Multi-task learning

## 1.2 Twitter における市民意見

## 1 はじめに

### 1.1 本研究の背景と目的

自治体による政策や計画の改善には、都市で暮らす市民の意見を反映させることが必要不可欠である。そのため、自治体は行政への市民参加に関する制度として、パブリックコメント制度や審議会委員による市民公募等を行っている [20] [21]。しかし、これらの既存の制度に対しては、「住民の意識・関心が低く、参加者が少ない」、「参加者の偏りや固定化などがある」といった問題が自治体によって指摘されており [20] [21]、偏りの無い、多くの市民意見を入手する新たな方法が必要と考える。

また、サービス業の経営戦略においても、顧客の声 (Voice of customers) は重要視されている [1]。さらに、他の市民が日頃各サービスに対して抱いている意見を参考にすることで、市民が利用するサービスを選択する際の助けになると考える。

このような背景から、本研究では、Twitter<sup>1</sup>における市民意見の自動抽出を目的とする。Twitter では多くのユーザが日頃感じたことを気軽に述べており、多様な市民意見を手軽に抽出できる。よって、意見の偏りや固定化を解決し、他の市民が日頃持っている意見を、市民が把握することに繋がると考える。

本研究では、日本で最も多くの市民が暮らす市である横浜市に着目し、新型コロナウイルス感染症の流行によって休園や登園自粛で大きな問題となった「保育園」に関する市民意見と、同じく新型コロナウイルス感染症の流行によって急激に利用者が増加した「飲食店のテイクアウトサービス」についての市民意見の抽出を試みる。

従来の多くの意見分析研究では、肯定や否定といった極性に着目して意見分析を行っている。また、Twitter のつぶやきを対象として、喜びや悲しみといった感情分類のアプローチで意見分析を行っている研究 [18] や、つぶやきから市民の要望を抽出する研究 [8] [9] も存在する。しかし、Twitter における市民意見は多岐にわたり、極性や感情の種類、要望といった特定の観点のみでは市民意見の分析には不十分であると考えられる。

そこで本研究では、複数の観点からつぶやきの分析を行うことで、Twitter における多様な市民意見を整理して抽出する。本研究における市民意見抽出では、深層学習モデルを用いて複数の観点からつぶやきを分析し、全ての未知のつぶやきに対して自動で観点の数だけ属性を付与する。そして、つぶやきに付与された複数の属性を組み合わせて条件を指定することで、抽出する市民意見を詳細に指定することが可能になる。

また、市民の意見はその時の社会の状況や自治体の対応によって推移していく [10] [14] [15] [18]。よって、本研究では時系列に着目して市民意見の出現頻度を分析するとともに、投稿時期を指定して市民意見を抽出することで、その時期の社会の状況や自治体の対応と関連のある市民意見を抽出する。

そして、本研究における市民意見抽出では、つぶやきの各属性の分類を同時に行う。この場合、共通のモデルを用いて複数のタスクを同時に行うことで、各タスクを独立して行うよりも高い精度が実現可能な、Multi-task learning [2] の有効性が期待できる。本研究では、複数の属性の分類において Multi-task learning を適用し、市民意見抽出の精度向上を目指す。

1: <https://twitter.com>

## 2 関連研究

### 2.1 アプレイザル理論を用いた意見分析に関する研究

アプレイザル理論 [13] は、選択体系機能言語学 (Systemic functional linguistics) [5] の考え方を基底とした理論であり、対象に着目した意見分析を行うことができる。

佐野ら [16] は、アプレイザル理論を用いた評価表現に基づいたブログの分類を行うことで、ブログの内容や分野によらず、ブログで表現される評価情報を一般的に分析できることを示した。著者の先行研究 [17] では、アプレイザル理論に基づき、コミュニティ QA 文書を用いた意見分析コーパスを人手で作成、分析することで、ドメインごとに出現する意見の特徴が異なることを示した。本研究では、アプレイザル理論による意見の種類を、人手ではなく、自動で推定することを試みる。

### 2.2 Twitter のつぶやきを対象とした意見分析に関する研究

栗原ら [9] は、ルールベース手法に基づき、つぶやきから自治体への要求を抽出した。また、川島ら [8] は、要求を含むつぶやきの抽出にサポートベクトルマシンを適用し、さらに学習データの収集に半教師あり学習の Distant Supervision を用いることで、低コストで高精度な要求抽出を実現した。本研究でも、市民の要求の抽出を試みるが、近年自然言語処理分野において、大規模なデータで事前学習済みの深層学習モデルを少数のデータによって各自の解決したいタスクにファインチューニングする手法の有効性が示されている。特に、Delvin らによって提案された BERT [3] は、感情分析を含む多くの自然言語処理タスクにおいて高い性能を示している。本研究では、BERT をファインチューニングすることでつぶやきを分析するアプローチによって、コストを抑えた意見分析を試みる。

著者らの先行研究 [6] [7] では、つぶやきに現れる市民意見を抽出し、意見の対象を明確にした上で 10 種類の意見分類を人手で行った。これによって、市民がどのような問題についてどのような意見を持っているのかを分析したが、人手による意見の種類の判断基準が曖昧という課題がある。本研究においても、深層学習モデルの教師データとして市民意見分析コーパスを作成する際、つぶやき内の意見の種類を人手で判定するが、明確な基準によって意見が体系化されているアプレイザル理論を用いることで、客観的なデータ作成を試みる。

### 2.3 時系列に着目して Twitter の市民意見を分析した研究

Kuan ら [10] は、Twitter における噂のライフサイクルを分析し、噂が広がるに連れて情報を含むつぶやきが増え、感情的なつぶやきが減少していくことを示した。Naskar ら [14] は、複数の話題についてのつぶやきを用いて、Twitter ユーザの感情の変化に影響を与える要因を分析し、リストのメンバーによるつぶやきがユーザの感情の変化に大きく影響することを示した。Pagolu ら [15] は、つぶやきに表れる意見の極性と株価の関係を分析し、意見の出現頻度が時期によって変化するだけでなく、株価の上昇、下降と関係していることを示した。

これらの研究から、Twitter におけるユーザの感情は社会の

状況やユーザの置かれた環境によって推移することが分かる。本研究でも、時系列に着目して市民意見を分析することで、市民意見の推移や出現頻度を分析する。

### 2.4 Multi-task learning に関する研究

Multi-task Learning [2] とは、共通のモデルを用いて関連する複数のタスクを同時に学習することで、各タスクを独立して学習するよりも高い精度を実現できる手法である。機械学習全般において、モデルが学習データに現れない未知のデータを予測するためには、未知のデータについての仮定、つまり帰納バイアス (inductive bias) が重要となる。Multi-task learning では、異なるタスクの学習によって得られた情報が帰納バイアスとしての役割を果たすことで、モデルの一般性を向上し、未知のデータに対する予測性能を高めることができる。

Liu ら [12] は、BERT を用いた Multi-task learning 手法である、Multi-Task Deep Neural Networks が、感情分析タスクを含む、自然言語理解の多くのタスクにおいて高い性能を発揮することを示した。また、Yadav ら [19] は、BERT の各層に重み付けを行うパラメータを、各タスク間のみではなく、さらに各層間でも共有することで、うつ症状検出タスクの精度向上を実現した。本研究では、意見に直接関連するタスクのみに留まらず、地域依存性等、つぶやきを複数の観点から分析するため、多くの自然言語理解タスクで高い性能を誇る Multi-Task Deep Neural Networks を参考に、BERT を用いた Multi-task learning モデルを構築する。

## 3 提案手法

本節では、複数の属性を用いた Twitter のつぶやきからの自動的市民意見抽出手法を提案する。

提案手法では、はじめにつぶやきに複数の属性を付与した市民意見分析コーパスを人手で作成する。本研究では、Twitter における横浜市民のつぶやきを用いて、「保育園」と「飲食店のテイクアウトサービス」の 2 つのドメインについて市民意見分析コーパスを作成する。続いて、作成した市民意見分析コーパスを教師データとして用い、各属性の分類モデルを訓練する。そして、訓練したモデルに大量の未知のつぶやきを入力し、属性を推定することで、全てのつぶやきに複数の属性を自動で付与する。これらの属性を任意の数だけ指定することで、未知の大量のつぶやきから、指定された条件を全て満たす市民意見のみを抽出することができる。また、属性のみでなく、つぶやきの投稿日によっても、抽出するつぶやきの条件を指定する。

そこで、3.1 節で提案手法においてつぶやきに付与する各属性を定義し、3.2 節で各属性の分類モデルについて説明する。

### 3.1 つぶやきに付与する属性の定義

本手法では、2 つのドメインを横断して付与する属性と、各ドメインにのみ付与する属性を定義する。各ドメインにのみ付与する属性は、特定の話題との関連性とする。各ドメインにおいて頻出すると考えられる特定の話題については、あらかじめ関連性を判断するモデルを訓練しておくことで、特定の話題に

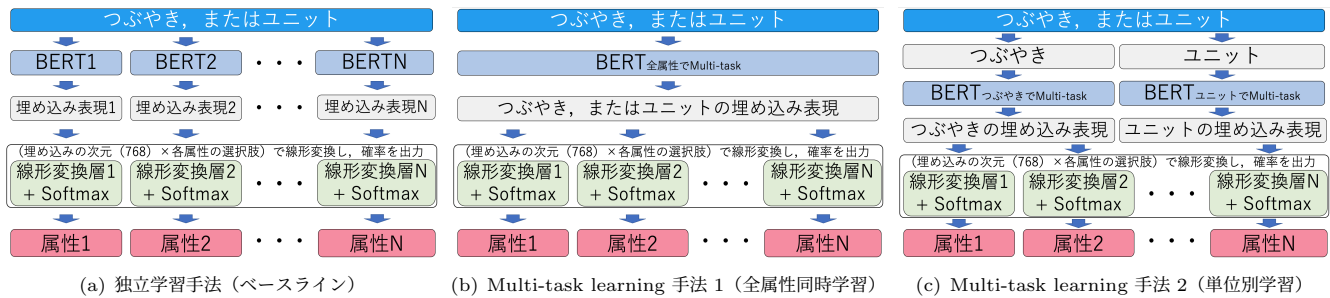


図 1 各属性の分類モデル

関連する意見を抽出できるようにする。つぶやきに付与する属性は以下の通りとする。

#### ドメインを横断して付与

##### ● 極性

行政や接客業の改善点を探す際には否定的な市民意見が重要となり、飲食店の高評価の口コミのような情報が欲しい際には肯定意見が重要となる。このように、市民意見を活用する場面によって必要な極性は異なるため、つぶやきの極性を判断する。選択肢は「肯定」、「否定」、「中立」、「意見無し」。

##### ● 意見タイプ

飲食店が接客を改善する際には店員の振舞に関する意見が重要となり、商品の改良をする際には商品の評価のような意見が重要となるように、意見の対象によって活用の方法は異なる。よって、アプレイザル理論に基づき、対象に着目した意見のタイプを判断する。

選択肢は「要求を表す意見」、「自発的感情の表明」、「人間・組織の振舞や行為を対象とした意見」、「事物・事象を対象とした意見」、「意見無し」。

##### ● 中立的な意見タイプ

著者の先行研究[17]を参考に、アプレイザル理論には表れないが、Twitterのつぶやきにおいて頻出する意見の種類を判断する。これによって意見の網羅性を高める。選択肢は「推測」、「提案」、「疑問」、「該当無し」。

##### ● 地域依存性

意見が都市で暮らす市民特有の意見か、もしくは社会一般的な意見かは、市民意見を社会に反映させる際に重要な要素となる。地域に依存するとは、市や県等の地名や、地域の特長が可能な飲食店等の施設名を含むものと定義する。選択肢は「依存」、「非依存」。

##### ● ドメインとの適合性

この属性は、単語としては各ドメインに関連するが、意味的には関連のないつぶやきからの意見抽出を避けるために定義する。たとえば、飲食店のテイクアウトサービスについての意見を抽出する際に、「仕事を家に持ち帰りすることになった」といった内容のつぶやきを「今日は美味しいお持ち帰り弁当を買った」といった内容と同様に抽出するのを避けることを目的とする。選択肢は「適合」、「不適合」。

##### ● 投稿主の立場

同じ内容についての意見であっても、市民の置かれた立場によって意見の方向性は異なる。よって、どのようなユーザによって

投稿されたつぶやきであるかを判断する。

選択肢は、保育園ドメインが「小さい子を持つ親」、「保育園関係者」、「その他」、飲食店のテイクアウトサービスドメインが「店を利用した人」、「飲食店」、「その他」。

#### 「保育園」ドメインにのみ付与

##### ● 休園・登園自粛との関連性

新型コロナウイルス感染症によって大きな問題となった、保育園の休園や登園自粛といった話題との関連性を判断する。

選択肢は「関連する」、「関連しない」。

##### ● 保育園の定員との関連性

横浜市内で大きな問題となっている待機児童問題や、保育園の可否のような保育園の定員の話題との関連性を判断する。

選択肢は「関連する」、「関連しない」。

#### 「飲食店のテイクアウトサービス」ドメインにのみ付与

##### ● 商品の評価との関連性

商品の味や量、提供状態等の評価を含むかを判断する。選択肢は「関連する」、「関連しない」。

つぶやきには複数の意見表現が現れることが多く、意見に直接関連する属性を付与する対象として、つぶやき全体という単位は大きすぎると考える。そこで、意見に直接関連する「極性」、「意見タイプ」、「中立的な意見タイプ」の3つの属性については、つぶやきをより細かく分割した、文、あるいは1つだけの意見が含まれる節を単位として属性を付与する（以降、文、あるいは1つだけの意見が含まれる節をユニットと呼ぶ）

### 3.2 各属性の分類モデル

作成した市民意見分析コーパスを用いて、各属性の分類モデルを構築する。本研究で用いるモデルの概要を図1に示す。

#### 3.2.1 独立学習手法（ベースライン）

本研究では、各属性の分類手法のベースラインとして、独立したBERTモデルによって各属性の分類を行う図1(a)の手法を用いる。この手法では、つぶやき、またはユニットを、属性の数だけ用意したBERTモデルへと独立に入力する。各BERTモデルから得られた埋め込み表現を各属性の選択肢の数へと線形変換し、Softmax関数へと入力することで属性を予測する。

#### 3.2.2 Multi-task learning 手法1（全属性同時学習）

Multi-task learningを用いた各属性の分類手法の1つ目として、図1(b)の単一のBERTモデルによって全ての属性を同時に学習する手法を用いる。この手法では、各つぶやき、ユニッ

トを全て単一の BERT モデルへと入力し、得られた埋め込み表現を各タスク固有の線形変換層へと入力する。そして、線形変換層の出力を Softmax 関数に入力し、各属性の予測確率を出力することで予測を行う。全ての属性の情報を単一の BERT モデルに集約することで、各属性が帰納バイアスとしての役割を果たし、モデルの一般性を高めることが期待できる。

3.2.3 Multi-task learning 手法 2 (単位別学習)

Multi-task learning を用いた各属性の分類手法の 2 つ目として、図 1(b) の 2 つの BERT モデルによってつぶやき単位、またユニット単位の属性をそれぞれ同時に学習する手法を用いる。この手法では、つぶやき、ユニットをそれぞれ別の BERT モデルへと入力することで異なる埋め込み表現を出力し、得られた埋め込み表現を各タスク固有の線形変換層へと入力する。そして、線形変換層の出力を Softmax 関数に入力し、各属性の予測確率を出力することで予測を行う。つぶやきとユニットを分割し、それぞれに複数の属性を付与する本手法においては、単位別に学習を行うことで 1 つの BERT モデルが学習するタスクの関連性を最大まで高める本手法は非常に有効であると考えられる。特に、アプレイザル理論を含むユニット単位の属性の同時学習を行う BERT モデルについては、意見に直接関連する属性のみを同時に学習するため、高い精度が期待できる。

全ての手法において、学習時に BERT は全 12 層をファインチューニングし、バッチサイズは 16 とする。また、Multi-task learning 手法においては、はじめに各タスクのデータを全てバッチへと変換し、得られた各バッチをランダムに並び替えた上でモデルへと入力する。

これらの手法のうち、各属性の分類に最も適した手法を明らかにすることで、精度の高い市民意見抽出を行うことを試みる。

4 実験：各属性の分類

本節では、提案手法を用いた各属性の分類実験について述べる。また、実験の詳細に先立って、本研究で使用するつぶやきの収集方法と、市民意見分析コーパスの各属性の人手によるアノテーション方法と結果について述べる。

4.1 市民意見分析コーパスの作成

4.1.1 横浜市民によるつぶやきの収集

はじめに、プロフィール情報に基づいて収集した横浜市民の Twitter アカウント 82,583 件のつぶやきを、Twitter の Streaming API を用いて収集した。このつぶやきのうち、2020 年 1 月 1 日から 2020 年 7 月 11 日までの計 28,971,414 件のつぶやきから、「保育園」、「保育士」、「保活」、「待機児童」の単語を含むつぶやきを保育園ドメインのつぶやきとして収集し、「持ち帰り」、「テイクアウト」の単語を含むつぶやきを飲食店のテイクアウトサービスドメインのつぶやきとして収集した。リツイートや重複するつぶやきは取り除いた。

本研究で用いるデータ数は各ドメインに共通で 2,622 件のつぶやきと、それらを文単位に分割したものとした。つぶやきの

表 1 各属性の判定者間一致度 (Fleiss の  $\kappa$  係数)

ドメイン	属性	一致度 (A チーム)	一致度 (B チーム)
保育園	地域依存性	0.921	0.900
	ドメインとの適合性	0.721	0.707
	休園・登園自粛との関連性	0.847	0.846
	保育園の定員との関連性	0.879	0.865
	投稿主の立場	0.841	0.830
	極性	0.673	0.693
	意見タイプ	0.642	0.635
テイクアウト	中立的な意見タイプ	0.745	0.764
	地域依存性	0.875	0.841
	ドメインとの適合性	0.913	0.835
	商品の評価との関連性	0.828	0.803
	投稿主の立場	0.841	0.830
	極性	0.778	0.766
	意見タイプ	0.722	0.781
	中立的な意見タイプ	0.783	0.754

文単位への分割には、Python のライブラリ、spaCy<sup>2</sup>を用いた。この際、名詞のみで構成される文は意見性を含まないものが多いため、spaCy による分割後に名詞のみで構成された文が生成された場合、つぶやきの先頭以外に現れる文は 1 つ前の文に結合し、先頭に現れる文は 1 つ後ろの文に結合する処理を行った。また、改行は文の区切りとした。さらに、ハッシュタグのみの文も意見を含まないものが多いため、前の文と結合する処理を行った。最後に、閉じ括弧から始まる文については、前の文と結合する処理を行った。

4.1.2 人手による各属性のアノテーション

収集したつぶやきに 3.1 節で定義した各属性を付与するアノテーション作業を人手で行い、意見分析コーパスを作成した。アノテーション作業は第一著者を含む合計 5 名の判定者によって行い、全てのアノテーション結果は多数決によって決定した。

アノテーション作業では、はじめに各判定者のアノテーション方針を一致させるための訓練を行った。訓練では、保育園ドメイン 250 文 (103 件のつぶやき) と飲食店のテイクアウトサービスドメイン 250 文 (134 件のつぶやき) の計 500 文 (237 件のつぶやき) を対象に、5 名全員で各属性の判定を行い、判定者のアノテーション方針が一致したところで訓練を終了した。この際、文単位に分割された文において複数の意見が含まれると判断した場合は、複数の意見が現れないユニットとなるまで分割した。この分割作業は全員で意見を交換し、過半数の判定者間で一致した文でのみ行った。

訓練終了後に、残りの全てのアノテーション作業を行った。第一著者は全てのつぶやき、文のアノテーションを行い、残りの 4 名を 2 名ずつに分けることで、各 3 名ずつの 2 チームを作り、各チームが半数ずつを担当した。この際も過半数の判定者間で一致した文については、複数の意見が現れないユニットとなるまで分割した。Fleiss の  $\kappa$  係数 [4] を用いた各チームのアノテーションの一致度を表 1 に示す。全ての属性において一致度が 0.6 (Substantial Agreement [11]) 以上となり、両チーム

2 : <https://spacy.io/>

の一致度も近い値であることから、判定者によって属性の判定に差異が生まれないことが示された。

アノテーション結果が 1-1-1 や 2-2-1 となり、多数決によって結果を決められない場合については判定者間で意見を交換することで、訓練データを含めた全ての結果を多数決で決定した。

最終的に得られた市民意見分析コーパスは、保育園ドメインが 2,622 件のつぶやきと 7,916 件のユニット、飲食店のテイクアウトサービスドメインが 2,622 件のつぶやきと 6,671 件のユニットとなった。

## 4.2 各属性の分類実験

### 4.2.1 目的

本実験では、3.2 節のモデルによる各属性の分類精度を評価し、各属性の分類に適したモデルを明らかにする。

### 4.2.2 方法

4.1 節で作成したコーパスの全てのデータについて、5 分割交差検証を用いて図 1 のモデルによる各属性の予測を行い、分類精度 (Accuracy) を確認する。5 分割交差検証の各検証では、全体の 20% のデータをテストデータとし、80% のデータを訓練データとする。また、訓練データのうち 20% を検証データとし、過学習を起こさない最適なエポック数を決定する。本研究では、つぶやきに現れる新語や固有表現に対応するため、新語辞書 (mecab-ipadic-NEologd) を用いた分かち書きが可能な、東北大学乾・鈴木研究室が公開している事前学習済み BERT モデル<sup>3</sup>を用いる。

しかし、飲食店のテイクアウトサービスドメインにおける「ドメインとの適合性」属性では 95% が適合、「中立的な意見タイプ」属性では 93% が意見無しと、データが極端に不均衡であり、データをそのまま用いて訓練を行ったところ全てのデータを多数派と予測してしまう結果となった。そこで、「ドメインとの適合性」属性では訓練データにおいて多数派データのダウンサンプリングを行い、「中立的な意見タイプ」属性では、保育園ドメインの中立的な意見タイプ属性における意見無し以外のデータ、つまり少数派データを用いて訓練データの拡張を行った。テストデータは、5 分割交差検証で作成したものをを用いる。

### 4.2.3 実験結果と考察

各属性の分類精度は表 2 の通りである。表 2 から、つぶやきを対象とした属性の推定については、各属性を独立して学習するベースラインにおいても概ね高精度で分類可能であることが分かる。飲食店のテイクアウトサービスドメインにおける商品の評価との関連性では Multi-task learning を用いることで分類精度が向上しているが、つぶやきを対象とした属性では大幅に精度が向上する属性は少ないことが分かる。

一方、ユニットを対象とした意見に直接関係する属性の推定については、ベースラインと比較して、Multi-task learning を用いる手法によって全ての属性で分類精度が向上している。特に、つぶやきとユニットを単位別に学習する手法では、分類精度が 10% 以上向上している保育園ドメインの極性をはじめ、大

表 2 各属性の分類精度 (Accuracy)

ドメイン	属性	ベース ライン	Multi-task	
			全属性	単位別
保育園	地域依存性	0.961	0.957	<b>0.968</b>
	ドメインとの適合性	<b>0.834</b>	0.808	0.819
	休園・登園自粛との関連性	0.924	<b>0.925</b>	0.923
	保育園の定員との関連性	0.959	0.953	<b>0.962</b>
	投稿主の立場	<b>0.740</b>	0.688	0.693
	極性	0.717	0.792	<b>0.825</b>
	意見タイプ	0.692	0.740	<b>0.767</b>
	中立的な意見タイプ	0.921	0.923	<b>0.930</b>
テイクアウト	地域依存性	<b>0.885</b>	0.871	0.868
	ドメインとの適合性	<b>0.962</b>	0.927	0.924
	商品の評価との関連性	0.905	<b>0.930</b>	0.891
	投稿主の立場	<b>0.814</b>	0.755	0.777
	極性	0.828	0.877	<b>0.890</b>
	意見タイプ	0.815	0.854	<b>0.873</b>
	中立的な意見タイプ	0.952	0.958	<b>0.960</b>

幅に精度が向上している。全属性を同時に学習する手法よりも、つぶやきとユニットを単位別に学習する手法の方が分類精度が高いことから、意見に関する属性を推定する際には、地域依存性やドメインとの適合性等のつぶやきを対象とした属性を同時に学習することで得られる情報はむしろノイズになっており、意見に関連するタスクで統一することで、各タスクの関連性を高めて Multi-task learning を行うことが有効であると考え。

また、両ドメインにおいて、投稿主の立場の分類精度が最も低くなっている。特に分類精度が低くなっている保育園ドメインの投稿主の立場において推定が不正解となっている例として、「保育士を休ませてあげてください」というように第三者目線から保育士のことを心配しているつぶやきを保育園関係者のつぶやきと推定している例や、他の保育園の対応について述べた後に「羨ましい」と一言述べている子を持つ親のつぶやきをその他と分類している等、つぶやきの細かな背景を捉えられていない多くの例が確認できた。投稿主の立場については、このような細かい背景を読み取り、テキストのみを用いて推定を行うには限界があると考える一方で、ユーザー名やプロフィールに「保育士」や「2 児の母」等の投稿主の立場が分かる情報を含むことがあるため、今後はこれらの情報を追加の特徴量として用いることで、分類精度の向上が期待できると考える。

## 5 実験：時系列に着目した市民意見の分析と未知のつぶやきからの市民意見抽出

### 5.1 目的

ドメインごと、また、意見タイプごとの出現頻度の推移を分析することで、市民意見の推移の傾向を確認する。そして、条件を変化させて未知のつぶやきから抽出した市民意見を確認することで、提案手法による市民意見抽出の有効性を検証する。

### 5.2 市民意見抽出の対象となるつぶやきの拡張

実験に先立って、作成した市民意見分析コーパスに含まれるつぶやきとは別に、4.1.1 節と同様の方法で新たに 2020 年 1 月 1 日から 2020 年 7 月 11 日のつぶやきを収集し、時系列分

3: <https://github.com/cl-tohoku/bert-japanese>



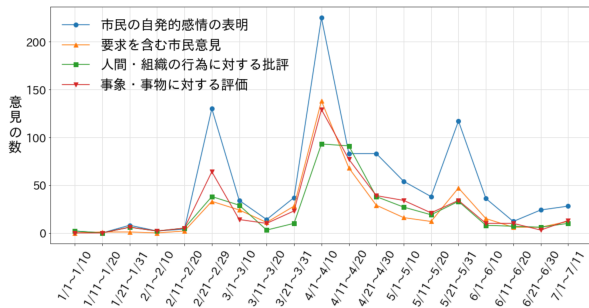


図2 保育園休園・登園自粛に関する意見数（推定）

析と市民意見抽出の対象となるつぶやきを拡張した。新たに収集したつぶやきは、保育園ドメインが6,913件、飲食店のテイクアウトサービスドメインが18,392件となった。文単位への分割についても4.1.1節と同様に行った。新たに収集したつぶやきから得られた文は、保育園ドメインが22,256件、飲食店のテイクアウトサービスドメインが54,772件となった。新たに収集したつぶやきについては、人手によるユニットへの分割は行っておらず、全て自動で前処理を行った。これらの未知のつぶやき、文を、作成した市民意見分析コーパスを教師データとして訓練した各属性の分類モデルに入力し、各属性を推定することで、自動で複数の属性を付与する。

### 5.3 方法

4.2.3節で得られた各意見の推定結果と、新たに収集したつぶやきについての推定結果について、出現頻度を10日間ずつの時系列順に並べ、時系列ごとの出現頻度の推移を分析する。今回抽出する意見は、保育園ドメインにおける休園・登園自粛に関連すると推定されたつぶやきの意見タイプと、飲食店のテイクアウトサービスドメインにおける意見タイプを対象とする。

本実験では、保育園ドメインの地域依存性属性の推定には単位別のMulti-task learning手法を用い、ドメインとの適合性属性の推定にはベースラインである独立学習手法を用いるといったように、各属性の推定には表2において最も分類精度が高いモデルを用いている。

提案手法を用いた未知のつぶやきからの市民意見抽出では、新たに収集したつぶやきと、4.2節の実験の5分割交差検証における各テストデータを未知のつぶやきとして行い、意見タイプや投稿時期を変化させることで、どのように異なる意見を抽出できるかを示す。

### 5.4 実験結果と考察

#### 5.4.1 市民意見の時系列推移の分析

保育園休園・登園自粛に関する時系列順の意見の出現頻度を図2、飲食店のテイクアウトサービスに関する正解データの時系列順の意見の出現頻度を図3に示す。

保育園休園・登園自粛に関する意見は、新型コロナウイルス感染症の影響で横浜市が小中学校の休校を発表した2月下旬、横浜市が緊急事態宣言発令後の保育園の休園に関する方針を発表した4月上旬、また横浜市が緊急事態宣言解除後の保育園の運営に関する方針を発表した5月下旬に急増している。飲食

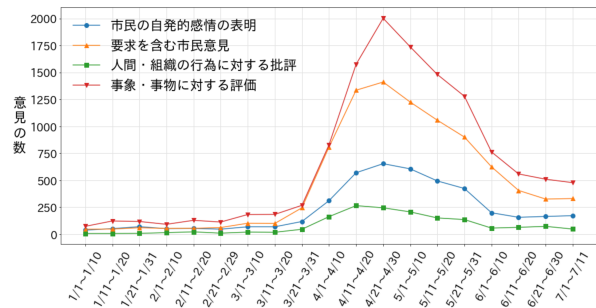


図3 テイクアウトサービスドメインの意見数（推定）

店のテイクアウトサービスに関する意見は、新型コロナウイルス感染症の流行が国内で本格化し、多くの飲食店がテイクアウトサービスを開始し始めた4月上旬に急増している。また、緊急事態宣言の解除により、多くの飲食店が通常営業を再開した5月下旬に意見が減少していることも分かる。このように、本手法により推定した市民意見の出現頻度は、社会情勢や社会情勢への自治体の対応と強く関係していることが分かった。

#### 5.4.2 意見タイプと投稿時期の変更による市民意見の変化

最後に、本研究の提案手法を用いた、大量の未知のつぶやきからの市民意見抽出において、条件を変更することで抽出可能な市民意見にどのような変化があるのかを確認する。本節では、意見タイプに着目して条件を変更した市民意見抽出を行った後に、投稿時期を変更して市民意見抽出を行う。

#### 1. 意見タイプの変更による抽出意見の変化

##### 保育園ドメイン

まず、保育園ドメインにおいては、図2で保育園の休園・登園自粛に関する市民意見が急増している4月上旬を投稿時期として指定する。4月上旬は、新型コロナウイルス感染症が拡大する中、横浜市が保育園の運営を継続すると発表した時期であり、この時期を指定することで市民の政策に対するフィードバックを抽出できると考える。抽出条件は以下の通りとする。

- (1) ドメインとの適合性：適合
- (2) 地域依存性：依存
- (3) 意見タイプ：自発的感情の表明
- (4) 休園・登園自粛との関連性：関連する
- (5) 投稿主の立場：小さい子を持つ親
- (6) 投稿時期：4月上旬

以上の条件で抽出された市民意見は以下の通りである。<sup>4</sup>

● 保育園休園しないとみんな会社行かなきゃいけないよね。保育園行かすの不安だな。。緊急事態宣言@横浜市

● 園児の保護者のコロナ発症が増えてきてる。

昨日入園した0歳児はただでさえ感染症をもらいやすいし、熱がコロナか風邪かの区別つかないだろうし不安しかない…横浜市、小学校も保育園も閉鎖して。

上記の条件を指定することで、横浜市に暮らす子育て中の親による、コロナ禍で子どもを保育園に通わせることに対する何かしらの感情表現を抽出できる。抽出結果から、横浜市の子育て中の親たちは、コロナ禍で子どもを保育園に通わせることについて不安感情を抱いている事が分かり、つまり自治体による政策に対して否定的な感情を抱いていることが読み取れる。

4：本稿に記載されている全てのつぶやきは表現を一部改変している。

続いて、新型コロナウイルス感染症拡大によって生じる不安を解決するためのアプローチの発見を試みる。そこで、不安を感じている市民が多く存在している4月の下旬のつぶやきから、条件を以下のように変更して市民意見を抽出する。

- (1) ドメインとの適合性：適合
- (2) 極性：肯定
- (3) 意見タイプ：人間・組織の振舞や行為を対象とした意見
- (4) 投稿主の立場：小さい子を持つ親
- (5) 投稿時期：4月上旬

以上の条件で抽出された市民意見は以下の通りである。

● うちもいよいよ保育園自粛のお願いきた！むしろ今まで来ていいと言っていたきありがとうございます…そして、市によると保育料は日割りで返ってくる 🙏  
これは助かる。

● 保育園側も保育士さんも感染対策を徹底してくれてる。園長先生は通勤中は帽子とメガネとマスクと手袋を着用して全部着替えてから仕事をするみたい 🙏  
保育士さんたちも家庭があるのにありがたいな…

意見タイプを人間・組織の振舞や行為を対象とした意見とし、極性を肯定とすることで、多くの親が不安を抱えている4月上旬に、保育園や行政はどのような対応をするべきなのかという、ベストプラクティスのような情報が抽出可能となっていることが分かる。このような情報は、新型コロナウイルス感染症のような未曾有の事態に直面した際に非常に有用であり、また本研究の自動的市民意見抽出によって、問題解決へのアプローチを見出すことができると考える。

#### 飲食店のテイクアウトサービスドメイン

続いて、飲食店のテイクアウトサービスドメインにおいて、まずは5.4.1節の時系列分析の結果をもとに投稿時期の指定を行う。図3から、飲食店のテイクアウトサービスに関する市民意見の数は、4月の下旬にピークを迎えていることが分かる。そこで、4月下旬のつぶやきから以下の条件で飲食店のテイクアウトサービスに関する市民意見の抽出を行う。

- (1) ドメインとの適合性：適合
- (2) 地域依存性：依存
- (3) 極性：肯定
- (4) 意見タイプ：事物・事象を対象とした意見
- (5) 商品の評価との関連性：関連する
- (6) 投稿主の立場：店を利用した人
- (7) 投稿時期：4月下旬

以上の条件で抽出された市民意見は以下の通りである。

● 近所のカニ屋のテイクアウトが悶絶するほどうまい!!  
正直カニクリームコロッケってあまり好きじゃないけど、このカニクリームコロッケはカニクリームコロッケの常識を覆すカニクリームコロッケ。とにかく美味しい!!!  
食べて！3個1500円!! #愛知屋 #横浜 #伊勢佐木町  
● 戸部のインドカレー シャンティデリ でテイクアウト。  
もともとデリバリーが得意なお店でメニューも充実。羊の挽肉入ナンのキーマクルチャが美味！ #横浜テイクアウト

上記の条件を指定することで、実際に飲食店のテイクアウトサービスを利用した人による高評価レビューのような情報が、未知の大量のつぶやきの中から自動で抽出できる。このような情報は、市民が利用する飲食店を選択する際に重要な意見となるだけでなく、飲食店側が提供した商品についてのフィードバックとして活用することも可能となっている。

さらに、飲食店のテイクアウトサービスドメインにおいて、意見タイプの条件を以下のように変更して市民意見抽出を行う。

- (1) ドメインとの適合性：適合
- (2) 地域依存性：依存
- (3) 極性：肯定
- (4) 意見タイプ：人間・組織の振舞や行為を対象とした意見
- (5) 投稿主の立場：店を利用した人
- (6) 投稿時期：4月下旬

以上の条件で抽出された市民意見は以下の通りである。

● 今日は関内弁天通りの #関内 kushiBa で、  
ベーコンオムレツと沖縄アグー豚のメンチカツ。  
マリノスの話をしてたら、“マリノス”と書いてくれました 🙏  
落ち着いたら飲みに行こう！  
● 米が浜通りの“鶏恵”さんがテイクアウトをやってる  
って聞いたので、特製つくねハンバーグ丼を買ってみました。  
オーダーしてから調理してくれるので出来たてです。  
サラダがセットにマスクまでついて税込み540円は  
お得ですよ(・∀・) #横須賀 #鶏恵

意見タイプを人間・組織の振舞や行為を対象とした意見とすることで、「オーダーしてから調理してくれる」等の、利用者が快く感じた店員の振る舞いに関する意見を抽出可能となる。このような意見は、飲食店が利用者の満足する接客を考える際に重要となる。特に新型コロナウイルス感染症のような未曾有の事態により、テイクアウトサービスといった店員にとって初めての試みをする際に非常に重要となると考える。

以上のように、アプレイザル理論に基づく意見タイプを変更して市民意見を抽出することで、市民意見を社会に反映させる際に、活用する場面に応じた意見を抽出することが可能となる。

#### 2. 投稿時期の変更による抽出意見の変化

最後に、時期による市民意見の推移を確認するため、投稿時期の条件のみを変更した以下の条件で市民意見を抽出する。

- (1) ドメインとの適合性：適合
- (2) 地域依存性：依存
- (3) 意見タイプ：自発的感情の表明
- (4) 休園・登園自粛との関連性：関連する
- (5) 投稿主の立場：小さい子を持つ親
- (6) 投稿時期：5月下旬

以上の条件で抽出された市民意見は以下の通りである。

● 横浜市から、緊急事態宣言解除後も6月末まで登園を  
自粛して欲しいと通知が届く。申し訳ないけどそれは難しい。  
経済活動再開、でも保育園に子どもは預けないで、は厳しい。  
● 横浜市も6月いっぱい保育園登園自粛だけど、  
もう限界なので少し預けるようにしよう。

テレワークで家でもやらなきゃいけないことが多くて、全然遊んであげられないのがほんとに申し訳ない。  
仕事のために観せたくもない動画サイトとか観させて、遊んでって言われても一人で遊んでって言うの辛い。

属性の指定方法は 5.4.2 節における 1 つ目の市民意見抽出と同様だが、投稿時期を 5 月下旬に指定している。5 月下旬は 4 月上旬と大きく社会の状況が異なっており、横浜市においては 6 月末までの保育園の登園自粛要請が発表された時期となっている。4 月の月上旬には保育園に子どもを通わせるのが不安だという意見が目立っていたにも関わらず、同じ条件でも 5 月の下旬にはむしろ早く保育園に子どもを通わせたいという市民意見が抽出されていることが分かる。これは、新型コロナウイルス感染症拡大当初は登園への不安が大きかったが、外出自粛によって働けない期間が続き、経済的な不安や精神的疲労が大きくなったためであると、抽出された市民意見から読み取れる。

このように、同じ条件で抽出された市民意見でも時期によって内容が異なるため、市民意見を適切に行政へと反映させるには時系列に着目した意見抽出を行う必要があると分かる。

## 6 おわりに

本研究では、Twitter のつぶやきを複数の属性を自動で付与し、属性と投稿時期によって条件を指定して市民意見を自動抽出する手法を提案した。また、各属性の関連性に着目した、Multi-task learning を用いることで一部の属性の分類精度が向上することを示した。

時系列に着目した意見分析においては、時系列ごとの市民意見の出現頻度から社会情勢や自治体の対応との関わりを読み取ることが可能であり、また、時期を指定して市民意見を抽出することによって、特定の社会情勢への自治体の対応と関わりがある市民意見を抽出可能であることを実験から示した。

今後の課題として、両ドメインにおける投稿主の立場の分類精度向上、意見抽出の対象となる地域の拡大が挙げられる。

## 謝 辞

本研究の一部は、科学研究費補助金基盤研究 B（課題番号 19H04420）の助成を受けて遂行された。

また、横浜市政策局共創推進課の関口昌幸さんには、行政視点からの貴重な助言を頂きました。ここに深く感謝します。

## 文 献

- [1] 青木保彦, 三田昌弘, 安藤紫. シックスシグマ. ダイヤモンド社, 1998, 220p.
- [2] Rich Caruana. Multitask Learning. Machine Learning. 1997, Vol. 28, p. 41–75.
- [3] Jacob Devlin, Ming-Wei Chang, Kenton Lee, Kristina Toutanova. “BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding”. 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers). Association for Computational Linguistics, 2019, p. 4171–4186.

- [4] Joseph L. Fleiss. Measuring Nominal Scale Agreement Among Many Raters. Psychological Bulletin. American Psychological Association, 1971, Vol. 76, No. 5, p. 378–382.
- [5] Michael Alexander Kirkwood Halliday. An Introduction to Functional Grammar. Edward Arnold, 1985, 387p.
- [6] 柏野和佳子, 立花幸子, 平本智弥, 関洋平. “市民意見の収集システムで得られたツイートからの「保育園」「教育」に関する意見抽出”. 言語処理学会第 23 回年次大会発表論文集. 言語処理学会, 2017, p. 533–536.
- [7] 柏野和佳子, 平本智弥, 関洋平. “市民意見の収集システムで得られたツイートからの道路・交通に関する意見抽出”. ことば工学研究会:人工知能学会第 2 種研究会ことば工学研究会資料. 2018, p. 13–21.
- [8] 川島崇秀, 佐藤哲司, 神門典子. “Twitter からの消費者ニーズの抽出手法に関する提案”. 第 8 回データ工学と情報マネジメントに関するフォーラム (DEIM 2016). 2016, B5-1.
- [9] 栗原理聡, 佐々木彬, 松田耕史. “Twitter を利用した地域毎の要望抽出”. 人工知能学会全国大会論文集. 一般社団法人人工知能学会, 2015, p. 1–4.
- [10] Chua Alton Yeow Kuan, Pal Anjan, Goh Dion Hoe-Lian. “Collective Sensemaking and Location-related Factors in the Context of a Brand-related Online Rumor”. The 22nd International Conference on Asian Digital Libraries. Springer International Publishing, 2020, p. 271–286.
- [11] J. Richard Landis, Gary G. Koch. The Measurement of Observer Agreement for Categorical Data. Biometrics. International Biometric Society, 1977, Vol. 33, No. 1, pp. 159–174.
- [12] Xiaodong Liu, Pengcheng He, Weizhu Chen, Jianfeng Gao. “Multi-Task Deep Neural Networks for Natural Language Understanding”. Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics. Association for Computational Linguistics, 2019, p. 4487–4496.
- [13] James Robert Martin, Peter Robert Rupert White. The Language of Evaluation: Appraisal in English. Palgrave Macmillan, 2005, 292p.
- [14] Debashis Naskar, Sanasam Ranbir Singh, Durgesh Kumar, Sukumar Nandi, Eva Onaindia de la Rivaherrera. Emotion Dynamics of Public Opinions on Twitter. ACM Transactions on Information Systems. 2020, Vol. 38, No. 2, 24p.
- [15] Venkata Sasank Pagolu, Kamal Nayan Reddy Challa, Ganapati Panda, Babita Majh. “Sentiment Analysis of Twitter Data for Predicting Stock Market Movements”. 2016 International Conference on Signal Processing, Communication, Power and Embedded System, 2016, p. 1345–1350.
- [16] 佐野大樹, 丸山岳彦. “評価表現に基づくブログ分類の試み—アブレザル理論を用いて—”. 言語処理学会第 16 回年次大会発表論文集. 言語処理学会, 2010, p. 174–177.
- [17] 関洋平. コミュニティ QA における意見分析のためのアノテーションに関する一検討. 自然言語処理. 2014, Vol. 21, No. 2, p. 271–299.
- [18] 鳥海不二夫, 榊剛史, 吉田光男. “ソーシャルメディアを用いた新型コロナウイルス禍における感情変化の分析”. 人工知能学会論文誌. 一般社団法人人工知能学会, 2020, Vol. 35, No. 4, p. F-K45-1-7.
- [19] Shweta Yadav, Jainish Chauhan, Joy Prakash Sain, Krishnaprasad Thirunarayan, Amit Sheth, Jeremiah Schumm. “Identifying Depressive Symptoms from Tweets: Figurative Language Enabled Multitask Learning Framework”. Proceedings of the 28th International Conference on Computational Linguistics, International Committee on Computational Linguistics, 2020, p. 696–709.
- [20] “市区町村における住民参加方策に関する調査研究 Part1”. 一般財団法人地方自治研究機構. [http://www.rilg.or.jp/htdocs/img/004/pdf/h24/h24\\_14\\_01.pdf](http://www.rilg.or.jp/htdocs/img/004/pdf/h24/h24_14_01.pdf), (参照 2020-12-12).
- [21] “市区町村における住民参加方策に関する調査研究 Part3”. 一般財団法人地方自治研究機構. [http://www.rilg.or.jp/htdocs/img/004/pdf/h24/h24\\_14\\_03.pdf](http://www.rilg.or.jp/htdocs/img/004/pdf/h24/h24_14_03.pdf), (参照 2020-12-12).