市民と地方議会の主張を比較するための属性ラベルの推定

妹尾 考† 関 洋平† 柏野和佳子†† 神門 典子††††,††††

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

E-mail: †s1811506@s.tsukuba.ac.jp, ††yohei@slis.tsukuba.ac.jp, †††waka@ninjal.ac.jp, †††kando@nii.ac.jp

あらまし 地方政治は地域に根差した政策が重要であり、民意が政治に反映されているかを確かめることが求められる。そこで、本研究では政治課題に対する市民と地方議会議員の立場(賛成・反対)との違いに着目し、双方の主張を比較するためのコーパスを作成した。政治課題は、「待機児童問題」、「カジノの誘致」の2種類とし、2つの都市を対象として市民と地方議会議員の主張を収集する。各主張には人手により、立場、有用性、地域依存性、適合性の4つの属性についてラベルを付与することで、具体的な根拠をもった主張や、市民の生活地域を対象とした主張に着目した分析を行う。また、コーパスを用いて事前学習モデルのBERTをファインチューニングすることで、ラベルの自動推定を試みる。

キーワード 地方議会会議録, Twitter, 立場分類, BERT, アテンション

1 はじめに

地方政治は地域に根差した政策が求められ、民意が政治に反映されているか確かめることが重要である。そこで、本研究では、地方政治における喫緊の課題や意見の割れている政策を政治課題と定義し、政治課題に対する地方議会議員と市民の立場(賛成・反対)の自動分類を行う。これにより大量のデータを分類して議員と市民の立場を比較する。

しかし、政治課題ごとに十分な量の訓練データを用意して分類モデルを構築することはコストを要する。また、対象とする文の中には、政治課題への明示的な賛否表現をもたず、立場の予測が困難な文も存在する。そこで、本研究では課題文へのアテンション機構 [10] を使用することで、立場の分類精度を向上させるとともに、単一の分類モデルによる政治課題を横断した学習の分類精度を改善し、訓練データ不足の解消を目指す。

本研究で対象とするデータは、地方議会については会議録から、市民については Twitter¹から収集する. 対象都市は横浜市と大阪市とし、政治課題は「待機児童問題」、「カジノを含む IR (統合型リゾート)の誘致」(以下、IR 誘致と呼ぶ)の 2 つを扱う. 各時期における、政治課題に対する賛否の割合の推移を調べることで量的な分析を行う. さらに、立場に加えて、(1) 根拠を伴っており議論に役立つか(有用性)、(2) 議員・市民の居住地域を対象とした主張であるか(地域依存性)、を判別することにより立場だけでは捉えられない質的な違いにも着目して議員と市民を比較する.

2 関連研究

2.1 課題を横断した立場分類と課題文の活用

Mohammad ら [9] が評価型ワークショップ SemEval-2016 で 提案した Task 6: Detecting Stance in Tweets のサブタスク Task B では、未知の課題に対するツイートの立場分類を行った. その結果、課題固有の訓練データが与えられる場合と比べて、分類精度が大きく低下することが明らかとなった. ゆえに、立場分類においては課題固有の訓練データが高い分類精度のために重要であるが、課題ごとに訓練データを十分な量用意することはコストを要する. そこで、本研究では、異なる政治課題を横断して学習することで、政治課題固有の訓練データ不足を補うことを図る.

しかし、Xuらは[11]は、従来の立場分類手法では、複数の課題を横断して単一の分類モデルで学習すると分類精度が低下することを示した。そして、課題を横断した学習の分類精度を改善するため、入力文だけでなく、課題文の埋め込み表現を考慮して立場を予測する LSTM ベースの分類モデルを提案した。本研究においても、事前学習モデルの BERT [3] をベースに、課題文へのアテンション機構を取り付けることで、課題文中の重要な語から情報を抽出して立場分類に活用する。これにより、2つの政治課題「待機児童問題」、「IR 誘致」を横断した学習の分類精度の向上を目指す。

2.2 立場以外の属性に着目した意見分析

石田ら [5] は、市民のもつ意見が多様であることに着目し、複数の属性に基づき市民のツイートを自動分類した. これにより、複数の属性を組み合わせることで市民の多様なツイートの中か

ら,条件を詳細に指定して意見を抽出した. さらに,市民の意 見数を時期別に集計することで,市民意見の傾向が社会情勢を 背景に推移していくことを明らかにした. 本研究においても, 議員と市民の多様な主張の中から,根拠を伴っており,議員・ 市民の居住地域を対象とした主張に着目して比較するため,有 用性と地域依存性も分類モデルにより判別する. 加えて,議員 や市民の立場の傾向は時期によって変化すると考える. そこで, 政治課題への賛否の割合を時系列に着目して分析する.

主張の立場以外の質の面に着目した研究には、Kimura ら [6] の事実検証可能性や、Kolhatkar ら [7] の Constructiveness がある。Kimura は、議員の発言の中から、客観的な情報をもつ根拠の伴った立場を抽出するため、立場に加えて事実検証可能性を分類した。本研究では議員だけでなく、個人的な情報を多く含む市民のツイートも対象とするため、客観性を重視する事実検証可能性では市民の有益な主張を捉えきれないと考える。

一方、Kolhatkar らはオンラインニュース記事に対する建設的なコメントを識別するため Constructiveness を定義した. Constructiveness とは単なる感情的な反応ではなく、必要に応じて適切な根拠を示したり、解決策や新しい視点・洞察を提供するなどして、議論を促進することを指す. 本研究では Kolhatkar らの Constructiveness と Kimura らの事実検証可能性を参考に、地方議会会議録と Twitter のツイート両方に付与可能な有用性を定義する.

3 提案手法

本提案手法は、まず地方議会会議録と Twitter から得たテキストに対し人手により各属性を付与することで、データセットを構築する.次にデータセットを用いて各属性の分類モデルを訓練する.続いて、訓練したモデルによって大量の未知のテキストの属性ラベルを予測して分析する.

そこで、3.1 節では提案手法で付与する属性について述べ、3.2 節で各属性を分類するためのモデルについて説明する.

3.1 属性の定義

本研究では以下の属性を人手によって付与したデータセット を作成し、深層学習モデルを訓練する.

● 立場

課題文に対する書き手の賛否を意味する. ラベルは「賛成」, 「反対」,「賛否両論」,「立場なし」である.

● 適合性

内容が政治課題と関連しているかを意味する. ラベルは「有り」、「無し」である. 政治課題「待機児童問題」の場合、保育所ではなく小学校の学童保育における待機児童は適合性「無し」となる. 政治課題「IR 誘致」の場合、IR とは関係のないカジノバーや、オンラインゲームのカジノの話は適合性「無し」となる.

● 有用性

内容が政治課題を巡る議論を促進し役立つものであるかを意味する. ラベルは「有り」、「無し」である. 有用性は Kimura ら

の事実検証可能性と、Kolhatkar らの Constructiveness を参考 に、表 1 のように条件を定める.

表 1: 有用性の条件

	- 数値,固有名詞,具体例等を含んだ具体的な根拠が
	あるもの
有り	- 問題点の指摘や新たな視点を提供するもの
	– 具体的なアイデアを提供するもの
	– 具体的な経験談を提供するもの
	- 根拠が曖昧なもの
無し	– 具体性に欠けるもの
	– 誹謗中傷・揶揄を含むもの

• 地域依存性

内容が対象地域(横浜市、大阪市)を対象しているかを意味する.ラベルは「有り」、「無し」である. 地名や施設名など地域特定が可能な語が含まれているものに加え、「待機児童心配だから市役所に相談してきた」など暗黙的に書き手の居住地域が対象であることが分かるものも地域依存性「有り」となる. 一方で、世界や日本全体を対象とした主張や一般論は地域依存性「無し」となる.

以上の4つが本研究で扱う属性である。なお、属性を付与する単位は、文単位と文書単位の2種類が存在する。文単位とは1文ごとに属性を付与することを意味する。文書単位とは地方議会会議録においては1発言中の1段落を指し、Twitterにおいては1ツイートを指す。地方議会における議論は基本的に一括質問一括答弁方式であり、発言者は1回の発言で複数の話題について述べることが多い。ゆえに、政治課題に対する立場を調べる際に、1発言全体を分析対象とするのは不適であると考える。地方議会会議録では発言は段落に分けて記述されているため、本研究では会議録の段落をTwitterのツイートに相当する単位として見なし、段落とツイートを文書として扱う。

属性を付与する単位について、石田ら [5] は 1 つのツイートには複数の異なる意見が含まれると考えて、文単位で属性を付与した。本研究においても文単位の属性付与を基本とし、文書単位の属性は各文の属性を踏まえて付与する。なお、地域依存性は文単位での判断は困難と考えて文書単位でのみ付与する。

3.2 分類モデル

本研究では政治課題を横断した学習における分類精度の向上を目指す。そのために、BERT に課題文へのアテンション機構を取り付けて、課題文中から重要な語から得られる情報を考慮して属性を分類する。これにより、分類モデルが政治課題の違いを区別できるようになり、分類精度が向上すると考える。加えて、課題文を考慮することで、政治課題への明示的な賛否表現をもたない文に対しても正しく立場の予測が行えることに期待できる。

本研究で用いる分類モデルの全体図を図1に示す。分類モデルは、3つのモジュールから構成される。文埋め込みモジュールでは、BERTと課題文へのアテンションを用いて、各文を埋め込み表現に変換し、文分類モジュールでは、埋め込み表現を用いて各文の属性を分類し、文書分類モジュールでは、各文の

埋め込み表現を合成することで文書の埋め込み表現を作成し, 文書の属性を分類する.

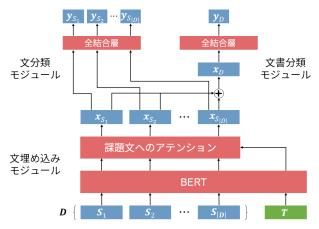


図 1: 分類モデルの全体図

3.2.1 文埋め込みモジュール

文埋め込みモジュールでは BERT により文の特徴を表す埋め込み表現を作成し、課題文へのアテンションにより課題文の情報を付与する.

BERT

文書 D は 1 文以上の文からなり,文 S により $D=\{S_1,...,S_{|D|}\}$ と表わされる.文 S_i はトークン t により $S_i=\{t_1,t_2,...,t_{|S|}\}$ と表わされる.文 S_i を BERT に入力することで,各トークンに対する埋め込み表現 $E_{S_i}=\{e_1,e_2,...,e_{|S|}\}\in\mathbb{R}^{|S|\times d}$ が得られる(通常,BERT では d=768 である).

なお、BERT は事前学習されたモデルを本研究で作成する データセットによりファインチューニングして用いる.

政治課題文へのアテンション

文書 D に対応する課題文 T は $T=\{t_1,t_2,...,t_{|T|}\}$ と表わされ,BERT に入力することで埋め込み表現 $E_T=\{e_1,e_2,...,e_{|T|}\}\in\mathbb{R}^{|T|\times d}$ が得られる.次に, E_{S_i} と E_T を全結合層で変換することで,クエリ $Q_i\in\mathbb{R}^{|S|\times d}$,キー $K\in\mathbb{R}^{|T|\times d}$,バリュー $V\in\mathbb{R}^{|T|\times d}$ を作成する.

$$\mathbf{Q}_i = \mathbf{W}_Q \mathbf{E}_{S_i} + \mathbf{b}_Q \tag{1}$$

$$K = W_K E_T + b_K \tag{2}$$

$$V = W_V E_T + b_V \tag{3}$$

続いて、 Q_i と K の内積をソフトマックス関数に入力することで、文 S_i の各トークンと課題文 T の各トークンとの関連度を表す A_i を計算する。 A_i と V の内積を E_{S_i} に加算することで、各トークンに課題文 T の情報を含んだ X_{S_i} が得られる。最後に、 X_{S_i} を平均プーリングに通して埋め込み表現の各次元で平均をとることで、課題文へのアテンションの出力 x_{S_i} が得られる。

$$\mathbf{A}_i = softmax(\frac{\mathbf{Q}_i \mathbf{K}^T}{\sqrt{d}}) \in \mathbb{R}^{|\mathbf{S}| \times |\mathbf{T}|}$$
 (4)

$$\boldsymbol{X}_{S_i} = \boldsymbol{E}_{S_i} + \boldsymbol{A}_i \boldsymbol{V} \in \mathbb{R}^{|\boldsymbol{S}| \times d}$$
 (5)

$$\boldsymbol{x}_{S_i} = average(\boldsymbol{X}_{S_i}) \in \mathbb{R}^d$$
 (6)

これにより、 x_{S_i} は文 S_i 中のトークンに加え、課題文 T 中の重要なトークンの情報を含んだベクトルになる.

なお,課題文へのアテンションは立場の分類精度向上を主たる目的とするが,適合性も政治課題と密接な関係をもつため,4章の評価実験では,全ての属性において課題文へのアテンションの有効性を検証する.

3.2.2 文分類モジュール

文単位の分類は、 x_{S_i} を全結合層で変換し、ソフトマックス 関数に入力することで行う.

$$\mathbf{y}_{S_i} = softmax(\mathbf{W}_S \mathbf{x}_{S_i} + \mathbf{b}_S) \in \mathbb{R}^c$$
 (7)

c は分類する属性のクラス数である.

3.2.3 文書分類モジュール

文書単位の分類は、文書 D の文書ベクトル x_D を、文書 D に 含まれる各文のベクトル $\{x_{S_1},...,x_{S_{|D|}}\}$ (これを $X_S \in \mathbb{R}^{|D| \times d}$ とする) の重み付き和によって作成し、 x_D を全結合層とソフトマックス関数に入力することで行う。各文の重みは文ベクトルを全結合層とソフトマックス関数で 1 次元に変換することで計算する。これにより、文書単位の属性を、文書に含まれる重要な文に着目して予測することができる。

$$\boldsymbol{w} = softmax(\boldsymbol{W}_w \boldsymbol{X}_S + \boldsymbol{b}_w) \in \mathbb{R}^{|\boldsymbol{D}|}$$
 (8)

$$\boldsymbol{x}_D = \boldsymbol{w} \boldsymbol{X}_S \in \mathbb{R}^d \tag{9}$$

$$\mathbf{y}_D = softmax(\mathbf{W}_D \mathbf{x}_D + \mathbf{b}_D) \in \mathbb{R}^c$$
 (10)

4 実験:課題文へのアテンションの有効性の検証

本章では、4.1節で実験に用いるデータセットの構築のためのデータの収集方法と人手によるアノテーションについて述べる. 次に 4.2 節で提案手法を用いて政治課題を個別に学習する実験について述べ、4.3 節で提案手法を用いて政治課題を横断して学習する実験について述べる.

4.1 実験データセットの構築

本研究で用いる実験データセットは2件の文書ジャンル(地方議会会議録,市民ツイート),2つの都市(横浜市,大阪市),2件の政治課題(待機児童問題,IR誘致)の組み合わせによる8種類のテキストから構成される.地方議会会議録は各種につき約1,000文,市民ツイートは各種につき1,000文から1,500文が含まれるように文書を抽出し,全体で約9,900文のデータセットとなる.

4.1.1 対象都市と政治課題

本研究は横浜市と大阪市を対象とする.以下に政治課題と, 政治課題を表す課題文を示す.本研究における立場とは課題文 に対する賛否を意味する.

- 待機児童問題:「行政の待機児童対策は成果が出ている, もしくは期待できる」
- IR 誘致:「カジノを含む IR(統合型リゾート)を誘致すべきである」

4.1.2 対象期間とクエリ

政治課題「待機児童問題」、「IR 誘致」について、2016年1月1日から2021月3月31日を対象に以下のクエリで収集する.

- 待機児童問題:「待機児童」,「保留児童」
- IR 誘致:「カジノ」,「IR」

4.1.3 地方議会会議録の収集

横浜市と大阪市の地方議会会議録²³から,定例会,臨時会,および各委員会の会議録を Web クローラーを用いて収集した。4.1.2 項で示した対象期間に合致する会議録は横浜市で 1,146件,大阪市で 809件である.これら会議録のうち政治課題のクエリを含む段落の中から,各都市,各政治課題でそれぞれ約1,000文が収集されるように段落を抽出した.段落の文への分割は句点「。」を基点に行った.

4.1.4 市民ツイートの収集

Twitter からプロフィール情報に基づき収集された横浜市民のアカウント 84,936 件と大阪市民のアカウント 56,307 件のツイートを Twitter の Streaming API を用いて収集した.これらのツイートのうち,4.1.2 項で示した対象期間中に投稿された横浜市民のツイート 121,015,645 件と大阪市民のツイート 30,396,883 件の中から,各都市,各政治課題でそれぞれ 1,000 文から 1,500 文が収集されるようにツイートを抽出した.なお,リツイートおよび重複するツイートは除外し,ツイート中のURL は削除した.ツイートの文への分割は,Python のライブラリ spaCy^4 を用いて行った.

また、Twitter には市民だけでなく議員や Bot、企業などのアカウントが存在する。そこで、非市民アカウントの条件を正規表現を用いて以下のように定めて、ツイートの収集対象から除外した。

議員アカウント

アカウント名またはプロフィール文が以下の条件を満たすもの. "[会院] 議員"にマッチし、かつ"応援|支持"にマッチしない

• Bot や企業などのアカウント

アカウント名が以下の正規表現にマッチするもの.

"(^|[^a-z])bot([^a-z]|\$)|ボット|新聞|news|ニュース|速報|公式|オフィシャル|official|宣伝|告知"

または、プロフィール文が以下の正規表現にマッチするもの.

"(公式|オフィシャル|official|宣伝|告知)[\s の用]?(アカ|垢|twi|ツイ|ツィ|account)"

4.1.5 人手による各属性のアノテーション

前述の方法で収集したデータに対し、3章で定義した属性をアノテーションすることでデータセットを構築する。アノテーションは著者および作業協力者6名の合計7名を3チームに分けて、分担して作業する。すなわち、協力者を2名ずつの3チームに分けて、著者は全てのチームに属することで3名からなるチームを作成する。これにより、各データにつき3名分のアノテーションが得られ、3名の多数決により最終的な値を決

定する. 作業者間で判断した属性ラベルが割れて多数決ができない場合, 作業者間で議論することでアノテーションを修正してラベルを決定する.

アノテーション作業は、はじめに訓練のためチーム間で共通の 1,000 文にアノテーションし、アノテーションの方針を一致させた、訓練の後、残りのデータを 3 チームで分担してアノテーションした。アノテーションの一致度として、各チームの Fleiss の κ 係数 [4] を表 2 に示す。全ての属性において値が 0.6 (Substantial Agreement [8]) を超えており、作業者間のアノテーションに差異が少ないことが示された。

表 2: 各属性の Fleiss の κ 係数

			κ係数		κ 係数		κ 係数	
文書ジャンル	政治課題	属性	(チーム 1)		(チーム 2)		(チーム 3 ¹)	
			文	文書	文	文書	文	文書
		立場	0.616	0.625	0.640	0.648	-	-
	待機児童	適合性	0.711	0.698	0.719	0.639	-	-
	問題	有用性	0.649	0.662	0.716	0.702	-	-
地方議会		地域依存性2	-	0.615	-	0.719	-	-
会議録	IR 誘致	立場	0.615	0.683	0.621	0.656	-	-
		適合性	0.649	0.799	0.637	0.679	-	-
		有用性	0.638	0.649	0.718	0.711	-	-
		地域依存性 2	-	0.794	-	0.725	-	-
		立場	0.617	0.674	0.614	0.662	0.632	0.690
	待機児童	適合性	0.661	0.748	0.692	0.656	0.793	0.936
	問題	有用性	0.646	0.708	0.661	0.709	0.751	0.843
市民ツイート		地域依存性2	-	0.773	-	0.686	-	0.818
印氏シィート	IR 誘致	立場	0.644	0.782	0.636	0.676	0.642	0.750
		適合性	0.693	0.919	0.763	0.827	0.786	0.920
		有用性	0.664	0.686	0.637	0.749	0.837	0.842
		地域依存性 2	-	0.745	-	0.778	-	0.816

¹チーム3は市民ツイートのみアノテーションした.

4.2 実験 1:政治課題を個別に学習

4.2.1 目 的

本節では、政治課題を個別に学習した場合において、課題文へのアテンションが有効に作用する政治課題、属性を確認する.

4.2.2 方 法

4.1 節で構築した実験データセットを用いて,5分割交差検証を行い,各属性の分類モデルを訓練・評価する.評価に使用する分類精度は文単位の分類と文書単位の分類それぞれのマクロ平均 F値の平均とする.分類モデルは文書ジャンル別,政治課題別に訓練する.なお,以下に示す属性は極端な不均衡データとなっており,予備実験では全てのデータを用いて学習した場合,少数派クラスの分類精度が非常に低かった.そのため,括弧で示す多数派クラスに属する文をダウンサンプリングし,50%削減した.

- 地方議会会議録:適合性(有り),立場(立場なし),地 域依存性(有り)
 - 市民ツイート:立場(立場なし),有用性(無し)

また,立場の「賛否両論」はデータセット中に 1%未満しか存在せず,多数派クラスをダウンサンプリングしても正確な分類は困難であったため,本実験では賛否両論の分類は行わないこととし、データセットから除去した.

比較手法として、提案手法から課題文へのアテンションを除

 $^{2 : {\}tt http://giji.city.yokohama.lg.jp/tenant/yokohama/pg/index.html}$

 $^{3 \ \}vdots \ https://ssp.kaigiroku.net/tenant/cityosaka/SpTop.html$

^{4:} https://spacy.io

 $^{^2}$ 3.1 節で述べた通り、地域依存性は文書単位でのみアノテーションを行ったため、文単位での κ 係数はない.

いた手法も同様に訓練・評価する。この手法の場合, 3.2.1 節 で述べた BERT の出力 $\mathbf{E}_{S_i} = \{\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_{|S|}\}$ が,文分類モジュールと文書分類モジュールへの入力となる.

4.2.3 結 果

各属性の分類精度を表3に示す.

表 3: 政治課題の個別学習の分類精度 (文と文書のマクロ平均 F 値の平均)

政治課	待機児童問題を		IR 誘致を		
以行詠	個別に学習		個別に学習		
課題文へのアテンション		使用	非使用	使用	非使用
	立場	0.696	0.697	0.667	0.700
地方議会会議録	適合性	0.653	0.675	0.619	0.618
地刀硪云云硪琢	有用性	0.849	0.841	0.770	0.764
	地域依存性	0.546	0.537	0.626	0.530
	立場	0.507	0.468	0.521	0.531
市民ツイート	適合性	0.726	0.676	0.846	0.843
中民ノイート	有用性	0.650	0.681	0.536	0.570
	地域依存性	0.857	0.856	0.882	0.882

4.2.4 考 察

立場については、市民ツイートの待機児童問題では課題文へのアテンションを使用した提案手法により分類精度が向上したが、他3つの場合においては精度向上がみられなかった。他の属性においては、適合性と地域依存性が3つの場合で分類精度が向上したが、有用性は市民ツイートでは分類精度が低下した。各政治課題を個別に学習した場合、訓練データは約2,000文しかないため、十分な訓練ができなかったと考える。したがって、次節の実験2では、政治課題を横断することで訓練データを拡張し、課題文へのアテンションの使用により分類精度が向上するか検証する。

4.3 実験 2:政治課題を横断した学習

4.3.1 目 的

本節では、2つの政治課題を横断した学習における課題文へのアテンションの有効性を検証する.

4.3.2 方 法

待機児童問題と IR 誘致の両政治課題のテキストが混交した訓練データを用いて、分類モデルを訓練する. 5 分割交差検証を行い、政治課題ごとに各属性の分類精度(文と文書のマクロ平均 F 値の平均)を求める. ダウンサンプリングの条件と賛否両論の処理は実験 1 と同様である.

4.3.3 結 果

各属性の分類精度を表 4 に示す. 立場に着目すると,提案手法により,市民ツイートでは待機児童問題で 8.2%, IR 誘致で 8%,分類精度が向上した. また,表 3 の政治課題の個別学習時と比較すると,両政治課題とも,政治課題を横断学習した場合,比較手法では分類精度が低下した一方で,提案手法では分類精度が向上している. このことから,政治課題の横断した学習においては,課題文を考慮することが重要であるとわかる.

4.3.4 考 察

地方議会会議録の待機児童問題において,提案手法と比較手 法の立場の分類結果を分析すると、「反対」意見は、提案手法が

表 4: 政治課題の横断学習の分類精度 (文と文書のマクロ平均 F 値の平均)

政治課題			政治課題を横断して学習				
			IR 誘致				
課題文へのアテンション		非使用	使用	非使用			
立場	0.698	0.691	0.683	0.674			
適合性	0.668	0.661	0.608	0.620			
有用性	0.850	0.853	0.772	0.769			
地域依存性	0.544	0.550	0.544	0.533			
立場	0.487	0.450	0.541	0.501			
適合性	0.721	0.693	0.840	0.837			
有用性	0.703	0.695	0.588	0.595			
地域依存性	0.869	0.848	0.882	0.885			
	デンション 立場 適合性 有用性 地域依存性 立場 適合性 有用性	規 デンション 使用 立場 0.698 適合性 0.668 有用性 0.850 地域依存性 0.544 立場 0.487 適合性 0.721 有用性 0.703	機児童問題	持機児童問題 IR デンション 使用 非使用 使用 立場 0.698 0.691 0.683 適合性 0.850 0.853 0.772 地域依存性 0.544 0.550 0.544 立場 0.487 0.450 0.541 適合性 0.721 0.693 0.840 有用性 0.703 0.695 0.588			

正解で比較手法が誤りのケースが 45 件であったのに対して,比較手法が正解で提案手法が誤りのケースは 11 件であった.提案手法だけが正解の場合,「疑問に思う」などの暗黙的に反対の立場を示す表現が捉えられていた. IR 誘致については,「反対」意見は,提案手法が正解で比較手法が誤りのケースが 36 件,比較手法が正解で提案手法が誤りのケースは 14 件であった.提案手法だけ正解の場合,言い回しがやや複雑なものが捉えられていたのに対して,比較手法だけが正解の場合は,「反対」など直接的な手がかり語が出現している事例がみられた.

市民ツイートの立場分類の「立場なし」の意見は、待機児童問題については、提案手法だけが正解のケースは 227 件であるのに対して、比較手法だけが正解のケースは 128 件であった。また、IR 誘致については、提案手法だけが正解のケースは 190件に対して、比較手法だけが正解のケースは 154件であった。一方「反対」意見は、待機児童問題では提案手法だけが正解のケースは 98件であり、IR 誘致では提案手法だけが正解のケースは 55件、比較手法だけが正解のケースは 55件、比較手法だけが正解のケースは 68件であった。すなわち、提案手法は市民ツイートの立場分類について「立場なし」の意見をより効果的に分類できているが、適合文に限ると、提案手法が比較手法のおよそ倍以上の件数を正しく分類できている。また、「ない」のような否定辞を含んでいる「立場なし」の適合文を正しく分類できている件数は、提案手法の方が多い。以上より、提案手法は課題文を考慮したうえでより適切に分類できるといえる。

また,市民ツイートの待機児童問題と IR 誘致について,提案手法により立場を正しく予測できた文の例を,課題文へのアテンションの可視化を併せて図 2 と図 3 に示す.

政治課題:待機児童問題 立場:賛成

課題文:[CLS]行政の<mark>待機</mark>児童対策は成果が出ている,もしくは期待できる[SEP]

[CLS]解決しないといけない問題はたくさんあるでしょうし、保育所中心の環境も見直す必要はあると思いますが、まずやれることを素早くやるのはいいと思います。[SEP]

[予測] 課題文へのアテンション 使用:賛成 不使用:立場なし

図 2: 立場の予測正解例(待機児童問題,市民ツイート)

政治課題:IR 誘致 立場:反対

課題文:[CLS]カジノを含むIR<mark>(</mark>統合リゾート)を誘致すべきである[SEP]

[CLS]市民の8割が誘致に反対なの分かってるからこその反応。[SEP]

[予測] 課題文へのアテンション 使用:反対 不使用:立場なし

図 3: 立場の予測正解例 (IR 誘致, 市民ツイート)

図2の文は、待機児童対策について賛同しているが、文中に「待機児童」の語はない.しかし、課題文中の「待機児童」に注意することで、立場を賛成と正確に予測している.また、対象文が待機児童対策の成果よりも施策自体へ理解を示していることは、課題文中の「期待できる」に注意していることと合致している.図3の文についても同様に、IR誘致に対して多くの市民が反対していることを訴えているが、文中には「IR」の語はない.しかし、課題文中の「IR」に注意することで、正確に立場を反対と予測している.

地方議会会議録については、提案手法により、待機児童問題で分類精度が 1%, IR 誘致で 1.3%分類精度が向上し、両政治課題ともわずかな精度向上となった。そこで、待機児童問題について不正解となった文の例を図 4 に示す。

政治課題:待機児童問題 立場:立場なし

課題文:[CLS]行政の待機児童対策は成果が出ている。もしくは期待できる[SEP]

[CLS]介護保険料を見ても、政令市や大阪府内で一番高い保険料は引き下げるべきです。[SEP] [予測] 課題文へのアテンション 使用:反対 不使用:立場なし

図 4: 立場の予測不正解例(待機児童問題,地方議会会議録)

図4の文は、介護保険料についての議論であり、待機児童とは直接関係ない.しかし、課題文へのアテンションを使用した場合、待機児童問題と結びつけて反対の立場と予測している.このように、課題文へのアテンションを使用することで、適合性のない文を賛成や反対と誤判定する傾向があることが分かった.したがって、今後は適合性を考慮した立場の予測が課題となる.解決策の1つとしてはマルチタスク学習[1]が有効と考える.マルチタスク学習により立場と適合性を共に学習させることで、適合性の学習結果から得たフィードバックによる立場の分類精度の向上を期待できる.

以上のとおり,立場分類については,課題文へのアテンションを使用することによる政治課題を横断した学習における一定の有効性を確認することができた.

5 分析:議員と市民の主張の比較

本節では、議員と市民の政治課題に対する主張について、まず5.1節で、時系列に着目した分析を行い、次に5.2節で有用性と地域依存性に着目した分析を行う.

5.1 分析 1: 時系列に基づく議員と市民の主張の分析

5.1.1 目 的

文書ジャンル別,都市別に政治課題に対する賛否の割合の推移を調べることで,議員と市民の主張の傾向を分析する.

5.1.2 方 法

提案手法を用いて各文書の属性のラベルを予測し、適合性「有り」、かつ地域依存性「有り」、かつ立場「賛成」または「反対」の文書を取り出す。文書は地方議会での発言日、または Twitter への投稿日に基づき、3ヶ月を1区間とする時系列順に並べ、同一の議員・市民によりグループ化して文書集合を作る。各文書集合内で、賛成の文書と反対の文書の数を集計し、多数派の立場がその時期の議員・市民の立場となる。

また、各時期において十分なデータ数を得るため、データセットの拡張を行った。4.1 節で述べた手法で再度データを収集し、表5に示す拡張データセットを作成した。

各時期で、各議員・市民の立場を求めた後、賛成派と反対派の割合を求めて折れ線グラフによって割合の推移を確かめる。グラフ中で賛否の割合が偏っている時期に焦点を当てて、その時期に含まれる文書の特徴を分析する。分析には単語単位のN-gram 頻度を用いる。賛成または反対の文書の中から、適合性が「有り」の文を対象に出現頻度を計上する。地方議会会議録ではN=4、市民ツイートではN=3とする。最後に、ワードクラウドを作成して主張の特徴語を可視化する。

表 5: 拡張データセットの統計情報

文書ジャンル	都市	政治課題	人数	*1文書数	*2 文数
	横浜市	待機児童問題	105	737	2090
地方議会会議録		IR 誘致	122	4658	14107
地刀俄云云皒邺	大阪市	待機児童問題	78	774	1652
		IR 誘致	109	2366	5587
	横浜市	待機児童問題	606	979	2748
市民ツイート		IR 誘致	4283	20615	60139
中氏ノイート	大阪市	待機児童問題	270	485	1360
		IR 誘致	1873	6745	20709

^{*1} 議会会議録は議員数,市民ツイートはアカウント数を指す
*2 議会会議録は段落数,市民ツイートはツイート数を指す

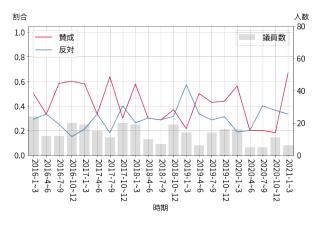
5.1.3 結 果

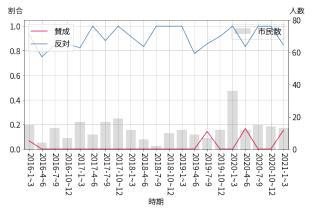
待機児童問題に対する議員・市民の立場の推移について,横浜市の結果を図5に示す. IR 誘致に対する議員・市民の立場の推移について,大阪市の結果を図6に示す. 各図では(a)が地方議会会議録,(b)が市民ツイートの立場を表している. なお,大阪市の待機児童問題と横浜市のIR 誘致の立場の推移の図については紙幅の都合上,掲載を割愛した. また,大阪市のIR 誘致についてのワードクラウドも割愛した.

5.1.4 考 察

待機児童問題

横浜市について、図 5(a) から、地方議会会議録では多くの時期で賛成が反対を上回っている。そこで、2016 年 7 月から2017 年 3 月に着目して N-gram 頻度によるワードクラウドを作成し、この時期の賛成派の特徴的な語を確かめる。結果を図7(a) に示す、「待機児童解消に向け」や「保育・教育コンシェルジュ」という語がみられる。保育・教育コンシェルジュは待機児童問題に対応するための横浜市の取り組みであり、この制度により待機児童解消を図っている様子がわかる。一方、図 5(b)から、市民ツイートでは全ての時期で反対派が多数を占めた、地方議会会議録と同様に 2016 年 7 月から 2017 年 3 月の反対派の市民ツイートについて、N-gram 頻度のワードクラウドを図 7(b) に示す、「保留児童」、「隠れ待機児童」などの語がみられる。これらから、横浜市が待機児童が少ない市と称される一方で、保留児童や隠れ待機児童などと呼ばれる待機児童に計上されない児童の存在を批判していることがわかる。

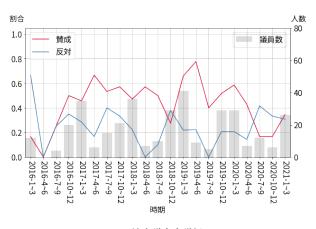


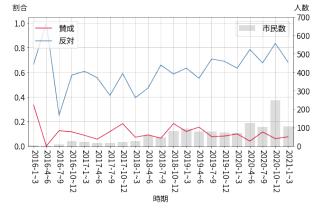


(a) 地方議会会議録

(b) 市民ツイート

図 5: 立場の推移(待機児童問題,横浜市)

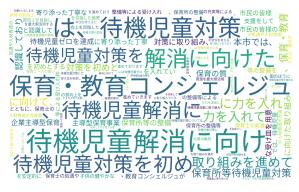




(a) 地方議会会議録

(b) 市民ツイート

図 6: 立場の推移 (IR 誘致, 大阪市)



(a) 地方議会会議録 (賛成, 2016 年 7 月から 2017 年 3 月)

(b) 市民ツイート (反対, 2016年7月から2017年3月)

図 7: 特徴語(待機児童問題,横浜市)

IR 誘致

大阪市では、図 6(a) から、地方議会会議録では賛成派が大半の時期で多数を占めている。2019年1月から9月の賛成派のワードクラウドには、「世界最高水準の成長」、「成長型 IR」、「国際観光拠点」などの語がみられ、大阪市が目指す IR の姿が捉えられた。一方で、図 6(b) より、市民ツイートでは全ての時期で反対派が多数であるが、どの時期にも賛成派が存在する。2019年1月から9月の賛成派のワードクラウドには、「国際会議」や「国際化」がみられるが、地方議会が「国際観光拠点」

を重視している点を考慮すると,地方議会と市民の賛成派が共通してIRに「国際」の観点から期待していることがわかる.

このように、時系列による立場の推移と、N-gram 頻度による特徴語の可視化により、議員と市民のそれぞれの主張の特徴を明らかにできた.

5.2 分析 2:有用性と地域依存性を用いた主張の抽出

5.2.1 目 的

有用性と地域依存性により、どのような主張が抽出可能か検証し、抽出された議員と市民の主張の特徴を比較する.

5.2.2 方 法

分析 1 で予測されたデータの中から有用性「有り」かつ地域 依存性「有り」と予測された文書を取り出して内容を確認する.

5.2.3 結果と考察

待機児童問題

横浜市の地方議会会議録と市民ツイートでは,表6の文書が 有用性「有り」,地域依存性「有り」の主張として抽出された.

表 6: 主張の抽出 (待機児童問題, 横浜市)

地方議会会議録 (賛成)

横浜平成 28 年 4 月 1 日現在の保育所等利用待機児童数は、認可保育所や小規模保育事業等の多様な保育施設の整備・拡充やきめ細かい相談支援サービスなどに取り組んだ結果、7 人となりました。

市民ツイート(反対)

そして、1 次で落ちたので育休延長したママは、おそらく利用を「取下げ」すると思う。だって、年度の途中で復職したところで、代替の人が入っているから、仕事はないんだもの。そういうママは待機児童ではなく「保留児童」でカウントされるんだよね。納得いかないよ

地方議会会議録では、具体的な取り組みとその成果である待機児童数を示しており、「横浜」から横浜市の話であると分かる。 一方、市民ツイートでは、市民の実情に沿って待機児童の不条理な点を指摘しており、問題を知るうえで有用といえる。

IR 誘致

大阪市の地方議会会議録と市民ツイートでは,表7の文書が 有用性「有り」,地域依存性「有り」の主張として抽出された.

表 7: 主張の抽出 (IR 誘致, 大阪市)

地方議会会議録 (賛成)

IR の立地を進める夢洲は負の遺産と言われてきましたが、いよいよ IR を核とする国際観光拠点として、その活用に向けた具体的な動きが本格化する段階に入りました。まさに大阪経済の起爆剤となるものであり、関西経済同友会の試算によると、経済効果は毎年 7,596 億円生まれ、雇用創出は $9~\mathrm{F}$ 8,000 人となっています。さらに、本市水道局による水需要の試算だけでも年間 5 億円となっており、我が会派として大いに期待するところであります。

市民ツイート(賛成)

南大阪の道路網を充実させるためにも、まず万博、IR をはじめとする 大阪市含めた大阪全体の成長を図り府の税収をあげ財源確保していく 必要があります。

地方議会会議録では、大阪市に誘致予定の IR の経済効果、 雇用創出効果に着目して IR 誘致に賛同している.一方、市民 ツイートでは、大阪市への IR 誘致がなぜ必要か、他の政策と の関係を考慮して主張している.

以上のように、有用性に着目することで、具体的な数値を用いた主張に加えて、市民の実情や IR 誘致の目的などが抽出された.こうした主張は議論を促進する有用な主張であり、政治課題をめぐる議論を把握するうえで抽出する価値がある.また、地域依存性に着目することで、国全体の話ではなく議員・市民の居住地域を対象とした主張を抽出できることを確認した.

6 おわりに

本研究では、地方議会会議録の議員の発言と Twitter の市民 のツイートに対して、適合性、立場、有用性、地域依存性の 4 つの属性を付与して,双方の主張を比較する手法を提案した. また,課題文へのアテンション機構を使用することで,多くの 場合において分類精度が向上することを示した.

時系列に基づき議員と市民の立場の推移を調べ、N-gram 頻度を用いて分析することで、議員と市民の主張の傾向に違いがある一方で、政治課題に対して議員と市民が共通して期待する要素も存在することを示した。さらに有用性と地域依存性に着目して議員と市民の主張を抽出することで、取り出す価値の高い主張や地域固有の主張が抽出可能であることを示した。

今後の課題は、マルチタスク学習による分類精度の向上や時系列分析への動的トピックモデル[2] の適用等があげられる.

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ました。また、西部みちる様には、分類モデルのエラー分析について多大なご協力をいただきました。心より御礼申し上げます。

文 献

- Rich Caruana. Multitask Learning. Machine Learning. 1997, Vol. 28, p. 41–75.
- [2] David M. Blei, John D. Lafferty. "Dynamic topic models". Proc. of the 23rd International Conf. on Machine Learning. ACM, 2006, p. 113–120.
- [3] Jacob Devlin, Ming-Wei Chang, Kenton Lee, Kristina Toutanova. "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding". Proc. of the 2019 Conf. of the North American Chapter of the ACL, Volume 1 (Long and Short Papers). ACL, 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] 石田哲也, 関洋平, 柏野和佳子, 神門典子. "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". 第 13 回デー タ工学と情報マネジメントに関するフォーラム (DEIM2021). 2021, D13-5.
- [6] Yasutomo Kimura, Hideyuki Shibuki, Hokuto Ototake et al. "Overview of the NTCIR-14 QA Lab-PoliInfo Task". The 14th NTCIR Conference on Evaluation of Information Access Technologies (NTCIR-14), 2019, p. 289-315.
- [7] Varada Kolhatkar, Maite Taboada. "Constructive Language in News Comments". Proc. of the First Workshop on Abusive Language Online. ACL, 2017, p. 11-17.
- [8] 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.
- [9] Saif Mohammad, Svetlana Kiritchenko, Parinaz Sobhani, Xiaodan Zhu, Colin Cherry. "SemEval-2016 Task 6: Detecting Stance in Tweets". Proc. of the 10th International Workshop on Semantic Evaluation (SemEval-2016). ACL, 2016, p. 31-41.
- [10] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, Illia Polosukhin. "Attention is All You Need". The 31st International Conference on Neural Information Processing Systems. Curran Associates Inc, 2017, p. 5998-6008.
- [11] Chang Xu, Cécile Paris, Surya Nepal, Ross Sparks. "Cross-Target Stance Classification with Self-Attention Networks". Proc. of the 56th Annual Meeting of the ACL (Volume 2: Short Papers). ACL, 2018, p. 778-783.