

# 災害ツイートを対象とした場所参照表現の抽出における 過去事例の利用とその災害種が及ぼす影響調査

六瀬聡宏<sup>1</sup> 内田理<sup>1</sup>

<sup>1</sup>東海大学情報理工学部情報メディア学科  
{trokuse, o-uchida}@tokai.ac.jp

## 概要

本研究では災害時に流通するツイートから場所を参照する表現の抽出について検討する。今回は機械学習モデルによる固有表現抽出(NER)の問題として解決を試みる上で1つの条件を設け、学習に過去の災害時に流通したツイートを利用することを考える。災害という状況下では、迅速かつ的確な対応が求められる一方で、多くのNERモデルは、性能の担保には正確にアノテーションされたデータが必要である。本研究では日本で発生した地震、豪雨、台風の3つの災害に焦点を当て、同じ種類の災害で学習した場合、異なる種類の災害で学習した場合の2つの観点から過去の災害時に流通したツイートの利用可能性を検証した。

## 1 はじめに

大規模災害時の被害を最小限に抑えるためには、迅速かつ的確な情報の収集と伝達が重要である。そのため、Twitterを始めとする即時性が高く利用者の多いソーシャルメディアの利活用に関心が寄せられている。災害時にはTwitter上で被災状況の発信や収集が盛んに行われており[1]、流通する情報は初動対応において重要な情報を含むことが報告されている[2]。このような背景から、災害対応にあたる政府や自治体でもTwitterの利活用が進められている[3]。

しかし、Twitterの災害時利用には解決すべき課題も残されている。そのひとつが流通する情報量の急激な増加である。例えば、2018年に発生した大阪府北部を震源とする地震の際には、発生直後の10分間で「地震」という単語を含むツイートが少なくとも27万件以上投稿されたとの報告がある[4]。

以上より、膨大なツイート群から、災害対応の意思決定に有益な情報のみを選別する必要がある。選別の観点として、信憑性や緊急性の他、各ツイートの内容がどの場所を対象としたものであるかを特定

することも重要である。ツイートには言及している地点の特定に有益な場所を参照するような表現(場所参照表現)を含む場合がある(表1)。場所参照表現が適切に抽出できれば、ツイートを地図にマッピングすることができ[5]、災害対応等により活用することが可能となる。

本研究ではツイート本文を対象に固有表現認識(NER)による場所参照表現の抽出を試みる。ただし、災害発生時にリアルタイムに適用することを想定し、現在発生している災害に対して、過去の災害時に流通したツイートで学習することを考える(図1)。多くの場合、NERは機械学習モデルによって実現され、対象のドメインのデータで学習されている。しかし、災害の発生と同時に学習に必要なデータの収集や選別などの一連の作業には多くの時間を要する。一方で、被害を最小限に抑えるという観点からは迅速な対応が必要不可欠である。これらの問題を解決する第一歩として、国内で発生した、豪雨、台風、地震の3種類の災害時のツイートからデータセットを構築し、過去の災害ツイートの利用可能性について検証する。今回は同じ種類の災害を利用した場合、異なる種類の災害を利用した場合で、モデルの性能にどのような影響を与えるかを調査した。

表1 場所参照表現を含む災害ツイートの例

Tweet1	@USERNAME 大雨の影響で安芸矢口駅付近の道路が冠水・浸水 現地の様子 <URL>
Tweet2	現在町内の河川はまだ大丈夫ですが、突然の増水に注意しましょう #九度山町

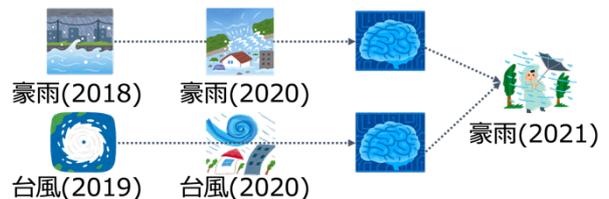


図1 過去の災害データを利用した学習の例

## 2 関連研究

災害時における Twitter 利用に関しては、対応の緊急度推定[6]や、キーフレーズ抽出[7]など幅広い分野で研究が進んでいる。これらの研究の多くは英語のツイートを対象としている場合が多いことから、他の言語でデータセットを構築し検証を進める動きもある[8][9]。また、これらのデータセットを統合して多言語に対応したモデルを構築する研究[10]も進められているが、多言語モデルは特定の言語に特化しているとの報告[11]もあることから英語以外の言語での検証の余地は残されている。

また、災害時におけるツイートの位置情報についても先行研究でいくつかの結果が示されている。例えば、ツイートが言及している地点の特定には、ツイートに付与されたジオタグと呼ばれる緯度経度情報が利用できる。しかし、ジオタグが付与されたツイート数は決めて少ないことが報告されているため[12]、他の方法と併用することが望ましい。このような背景からツイート本文を対象とした NER によるアプローチは数多く存在し、洪水[13]や地震[14]といった状況下での研究がある。特に、Suwaileh ら[15]はラベル付きのツイートが利用できない場合に、直近で発生した災害のツイートの有効性について検証している。この Suwaileh らの研究は本研究と立ち位置が非常に近いが、日本国内の災害時ツイートを対象とした場合、考慮すべき点が存在する。まず、災害の発生頻度や規模、その種類は地域によって異なる。例えば、日本国外では砂嵐 [16]や山火事[17]に関する研究などが進められているが、日本国内ではこれらの発生頻度や規模は小さい。その一方、日本では地震が発生する頻度や規模が他国に比べて大きく、東日本大震災や熊本地震など、甚大な被害をもたらしたケースが少なくない。次に、英語と日本語による言語の性質の違いが挙げられる。日本語は単語間の境界が曖昧であり、単語への分割方法が性能に影響を与えることが指摘されている[18]。また、文法の違いも考慮すべき要素である [19]。

従って、日本国内で発生した災害に対し、日本語で記述されたツイートに対する検証の余地はまだまだ大きいと考える。

真備	の	地域	交流	センター	が	停電
U	O	B	I	L	O	O

図 2 問題設定の概要

## 3 問題設定

本研究における場所参照表現の抽出とは、ツイートに含まれる全ての場所参照表現を特定することである。問題設定の概要を図 2 に示す。

$L$ 個の単語 $w_i (i = 1, \dots, L)$ から構成されるツイート $t = \{w_1, w_2, \dots, w_L\}$ が与えられたとき、 $t$ に含まれる場所を参照している単語 $w_i$ をすべて特定することを考える。場所参照表現は複数の単語にまたがる可能性があるため、ツイートを単語の系列とみなし、各単語のクラスを予測する系列ラベリング問題として定義する。クラスの定義は Yang ら[20]を参考に BILOU の 5 クラスを利用する。“B”は場所参照表現の先頭となる単語を示し、次の単語も場所参照表現の場合は“T”を割り当てる。末端まで到達した場合に“L”を割り当てることにより、“B”から“L”までが場所参照表現であることを表現する。場所参照表現ではない単語は“O”を割り当て、1 単語からなる場所参照表現は“U”を割り当てることで表現する。本研究ではこの定義に従い、ツイートを構成する単語列に対してマルチクラス分類を試みる。

## 4 データセット

### 4.1 ツイートの収集

日本国内で発生した豪雨、台風、地震の 3 種類の災害に焦点を当て、検証に利用するツイートを収集する。今回は Twitter Search API を利用し付録 A に示す災害時に流通したツイートを収集した。本研究では、以下の時間を起点として、24 時間以内に流通したツイートを対象とした。

- 地震：地震発生時刻
- 豪雨：気象庁から最初に大雨特別警報が発表された時刻から 6 時間前
- 台風：気象庁から日本国内に上陸の発表があった時刻から 6 時間前

災害に関係ないツイートを除外するため、Paul ら[21]などを参考に、事前の予備実験に基づいて決定したクエリを指定する。詳細は付録 A に示す。

### 4.2 災害に言及したツイートの絞り込み

Paul ら[21]の調査によると、災害に関係するクエリを含むツイートでも実際に災害について言及しているツイートは非常に少ないことが知られている。本研究では更に以下の条件をすべて満たすオリジナ

ルツイート（リツイート以外のツイート）のみに絞り込みを行う。

- (1) 日本語で記述されている
- (2) 1回以上リツイートされている
- (3) Twitter 公式クライアントから投稿されている
- (4) 1つ以上の場所参照表現を含んでいる

これらの条件は、一度でも拡散されたツイートは有益な情報を含み、かつ、現地からの投稿は機械的に投稿する BOT とは異なり、特別な Twitter クライアントを利用していないことがほとんどであろうという仮説に基づいて設定した。条件(4)に関しては関連研究[13]などを参考にしてオープンデータを利用した辞書を構築し、最長一致で文字列マッチングを行うことにより条件の充足を判断した。それ以外の条件に関してはツイートのメタデータを参照して機械的に判定を行った。

### 4.3 アノテーション

前節の処理で絞り込まれたツイート群に対して、含まれる場所参照表現にアノテーションを行う。今回は Matsuda ら[22]の研究で提案されているアノテーションスキームを参考に実施した。対象となった場所参照表現の例を付録 C に示す。作業の過程で“球磨川”の様な場所ではなく人名を指している場合や、実際の被害状況ではなく被災地の状況を心配するようなツイートを除外し、各災害に必要なツイート数が得られるまで作業を繰り返す。以上の手順を踏んで、豪雨、地震、台風の3種類の災害に対し、それぞれ3つの事例を準備した。今回は災害ごとに1,000ツイートを用意し、合計9,000ツイートのデータセットを構築した。

なお、データの偏りといった観点からアノテーション対象のツイートはランダムサンプリングによる選択が望ましい。しかし、今回は災害に関連したクエリを含むツイートだけでも膨大な数のツイートが流通しているため、前述の条件で絞り込みを行った。詳細を付録 A に示す。また、作成したデータセットは再現性や研究コミュニティの活性化の観点から公開されることが望ましい。しかし、災害ツイートに含まれる場所参照表現は Twitter ユーザの居住地や所属組織などプライバシーと密接に関係しているため、慎重になるべきであると考えた。その一例として、救助を求めるツイートに個人の住所が記載され

ているケース[23]を複数確認している。このような背景から今回はデータセットの公開を見送り、構築方法のみを示すこととした。

## 5 実験

### 5.1 実験設定

本研究では場所参照表現の抽出に BERT[24]を採用した。BERT は系列ラベリング問題のベースラインとして広く利用されており、比較的少量の学習でも良好な結果を得られることが知られている。また、BERT のような大量のデータで事前に学習を行う基盤モデルには ALBERT[25]や RoBERTa[26]が存在するが、Suwaileh ら[15]の研究との整合性を考慮した。前述のような基盤モデルがより高い性能を発揮するためには、対象のドメインのコーパスから学習することが望ましい[27][28]。しかし、学習に必要なコストが高いため、事前学習済みの BERT モデル (<https://huggingface.co/cl-tohoku/bert-base-japanese-v2>) を利用しデータセットで微調整する。

また、学習の効率化のため、ツイート中の URL やユーザ名は<URL>と<USERNAME>にそれぞれ置換した。学習時のハイパーパラメータはオリジナルの論文[24]を参考に optimizer として AdamW を選択し、learning rate を  $2e-5$ 、batch size を 8 に設定した上で epoch を 5 として学習を行った。

以上の設定に基づき、過去の災害ツイートの利用を2つの観点から検証する。まずは、同じ種類の災害で検証した場合を考える。例えば、令和3年8月豪雨での利用を考えた場合、それ以前に発生した2つの豪雨災害で学習し、令和3年8月豪雨のツイートで検証する。また、令和3年8月豪雨のツイートで学習した場合も検証する。この検証の目的は、発生中の災害で即時に学習データが利用可能な場合を性能の上限値とみなし、過去の災害で学習した場合と比較することである。さらに、災害の種類を変えて検証した場合の影響を検証する。同じく令和3年8月豪雨の利用を考えた場合、地震や台風で学習する場合を考える。比較として、同じ種類の過去の豪雨災害を利用した場合、発生中の災害で即時に学習データが利用可能な場合場合も併せて検証する。以上の実験にあたり、各災害のアノテーション済みツイートを Train(80%)、Dev(10%)、Test (10%)の3つの用途別に分割する。Train で学習し Test を推論し

た場合の性能を検証し、データの観察が必要な場合は Dev を利用する。

## 5.2 実験結果（同じ種類の災害）

同じ種類の災害で検証した場合の検証結果を付録 D に示す。同じ種類の災害で学習した場合（実験 ID 2, 3, 6, 7, 10, 11）でも、F 値で最大 6 ポイント以上の差を確認した。Dev で推論した結果の傾向を調査したところ、ピーク時の時間帯が影響を与えている可能性があることがわかった。例えば、日中に大雨特別警報が発令された平成 30 年台風 21 号の場合は、現地から具体的な災害の状況を確認できる有益なツイートを多く確認できた。一方で、令和元年房総台風の場合は、明け方の午前 3 時に接近し、午前 5 時に日本へ上陸したため、災害に言及するツイートは少ない。場所参照表現のバリエーションも限定されており、汎化性能に影響を与えたと考えられる。

以上の結果からデータの偏りを抑えるため、2 つの災害時に流通したツイートを混ぜ合わせた設定でも検証した（実験 ID 4, 8, 12）。学習データの数を合わせるため、2 つの災害から半分ずつランダムサンプリングし、学習と開発用のデータセットを構築した。実験の結果、豪雨と台風の場合で性能の向上を確認できた一方で、地震の場合は大きく性能が低下した。地震の場合は被害が広範囲に発生するため、ツイートに含まれる場所参照表現の粒度が大きい。そのため、含まれる場所参照表現が特定の都道府県名に偏っており、影響を及ぼしたと考えられる（実験 ID 8）。一方で、地震の場合は過去の地震を引き合いに出して注意喚起や被災報告するようなツイートも多く、場所参照表現が重複している傾向がある。この性質はプラスの影響を与える場合もあると考えられる（実験 ID 5）。

最後に、発生中の災害で即時に学習データが利用可能な場合も検討した（実験 ID 1, 5, 9）。地震は比較的小さいが、豪雨と台風の場合、性能には F 値で 3 ポイント程度の開きがあることもわかった。

## 5.3 実験結果（異なる種類の災害）

前節の検証結果から、複数の災害データを混合させることによって、汎化性能の向上が期待できる。その一方で、災害種によって組み合わせには検討の余地があることを示唆している。発生中の災害で即時に学習データが利用可能な場合と同等かそれ以上の性能を獲得するため、さらなる検証を行った。結

果を付録 E に示す。

5.2 での検証を踏まえ、災害の種類ごとに過去に発生した 2 つの災害を半分ずつ混ぜたツイートデータで学習を行った。付録 E に示す通り、豪雨に対し台風で学習した場合は、地震で学習した場合と比べると性能が高い。これは、台風は豪雨と同じく土砂災害や水害による被害が多く、ツイート内容の傾向が類似しているからであると考えられる。一方、台風に対して豪雨と地震で学習した場合を比較すると、地震の方が高い性能を示している。このことから、災害の傾向が似ている場合は性能向上が期待できるといった仮説には慎重になるべきであると考えられる。加えて、地震の場合は地震で学習するよりも、台風で学習した場合の方が良い性能を示している。交通インフラへのダメージなど、被害については共通する部分も少なからずあるが、風水害が被害の中心であり災害の性質が異なる場合でも良好な結果が得られたことは興味深い。これは災害の種類が異なるデータも学習に利用することが可能であることを示唆している。また、地震は単独での学習が適しているなど豪雨や台風とは異なる傾向を示しており、今後さらなる検証が必要である。

## 5.4 考察

これらの結果から、過去の災害ツイートと、発生中の災害ツイートを利用する場合では性能に開きがある。性能の向上にはマルチタスク学習や、CRF の利用が採用されることが多い。ただし、Wikipedia で学習された BERT はツイートの埋め込みに対して効果が低いといった報告もあり[29]、ツイートの学習したモデル[30]への変更は検討の価値がある。

最後に、先行研究である Suwaileh ら[15]との比較について述べる。台風と豪雨のような傾向の似た災害同士で学習が効果的であることなど、定量的な検証結果は概ね一致している。しかし、今回は定性的な分析まで踏み込んでいない。性能向上の観点からもエラーの傾向などが一致するかは、重要かつ興味深い疑問であり、今後の検討課題である。

## 6 まとめ

本研究では、災害時に流通したツイートからの場所参照表現抽出における、過去の災害ツイートの利用可能性を災害種の観点から検証した。今後は大雪などの災害種の拡張や、考察での検討事項の他、定性的な分析を進めることで性能の向上を試みる。

## 謝辞

本研究は、科研費基盤研究(C) 22K12277 の助成を受けて実施した。

## 参考文献

- [1] Saleem, H., Zamal, F., Ruths, D.: Tackling the Challenges of Situational Awareness Extraction in Twitter with an Adaptive Approach. *Procedia Engineering*, 107, pp. 301–311 (2015).
- [2] Meier, P.: Digital Humanitarians: How Big Data Is Changing the Face of Humanitarian Response. (2015).
- [3] Uchida, O., Utsu, K.: Utilization of Social Media at the Time of Disaster. *IEICE ESS Fundamentals Rev.* pp. 13, 301–311 (2020).
- [4] Yamada, S., Utsu, K., Uchida, O.: An Analysis of Tweets During the 2018 Osaka North Earthquake in Japan -A Brief Report. In: *Proc. 2018 5th International Conference on Information and Communication Technologies for Disaster Management (ICT-DM)*, pp. 1–5 (2018).
- [5] Uchida, O., Rokuse, T., Tomita, Kajita, Y., Yamamoto, Y., Toriumi, F., Semaan, B., Robertson, S., Miller, M.: Classification and Mapping of Disaster Relevant Tweets for Providing Useful Information for Victims During Disasters. *IEEE Trans. Image Electronics and Visual Computing*, 3, pp. 224–232 (2015).
- [6] Sarioglu Kayi, E., Nan, L., Qu, B., Diab, M., McKeown, K.: Detecting Urgency Status of Crisis Tweets: A Transfer Learning Approach for Low Resource Languages. In: *Proc. 28th International Conference on Computational Linguistics*, pp. 4693–4703. International Committee on Computational Linguistics, Barcelona, Spain (Online) (2020).
- [7] Ray Chowdhury, J., Caragea, C., Caragea, D.: Keyphrase Extraction from Disaster-related Tweets. In: *The World Wide Web Conference*, pp. 1555–1566. Association for Computing Machinery, New York, NY, USA (2019).
- [8] Alharbi, A., Lee, M.: Kawarath: An Arabic Twitter Corpus for Crisis Events. In: *Proc. 6th Arabic Natural Language Processing Workshop*, pp. 42–52. Association for Computational Linguistics, Kyiv, Ukraine (Virtual) (2021).
- [9] Cobo, A., Parra, D., Navón, J.: Identifying Relevant Messages in a Twitter-based Citizen Channel for Natural Disaster Situations. In: *Proc. 24th International Conference on World Wide Web*, pp. 1189–1194. Association for Computing Machinery, New York, NY, USA (2015).
- [10] Ray Chowdhury, J., Caragea, C., Caragea, D.: Cross-Lingual Disaster-related Multi-label Tweet Classification with Manifold Mixup. In: *Proc. 58th Annual Meeting of the Association for Computational Linguistics: Student Research Work-shop*, pp. 292–298. Association for Computational Linguistics, Online (2020)
- [11] Martínez-García, A., Badia, T., Barnes, J.: Evaluating Morphological Typology in Zero-shot Cross-lingual Transfer. In: *Proc. 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, pp. 3136–3153. Association for Computational Linguistics, Online (2021).
- [12] Cheng, Z., Caverlee, J., Lee, K.: You are Where You Tweet: A Content-based Approach to Geo-locating Twitter Users. In: *Proc. 19th ACM international conference on Information and knowledge management*, pp. 759–768. Association for Computing Machinery, New York, NY, USA (2010).
- [13] Al-Olimat, H., Thirunarayan, K., Shalin, V., Sheth, A.: Location Name Extraction from Targeted Text Streams using Gazetteer-based Statistical Language Models. In: *Proc. 27th International Conference on Computational Linguistics*, pp. 1986–1997. Association for Computational Linguistics, Santa Fe, New Mexico, USA (2018).
- [14] Middleton, S.E., Middleton, L., Modafferi, S.: Real-Time Crisis Mapping of Natural Disasters Using Social Media. *IEEE Intelligent Systems*, 29, 9–17 (2014).
- [15] Suwaileh, R., Imran, M., Elsayed, T., Sajjad, H.: Are We Ready for this Disaster? Towards Location Mention Recognition from Crisis Tweets. In: *Proc. 28th International Conference on Computational Linguistics*, pp. 6252–6263. International Committee on Computational Linguistics, Barcelona, Spain (Online) (2020).
- [16] Hamoui, B., Mars, M., Almotairi, K.: FloDusTA: Saudi Tweets Dataset for Flood, Dust Storm, and Traffic Accident Events. In: *Proc. 2th Language Resources and Evaluation Conference*, pp. 1391–1396. European Language Resources Association, Marseille, France (2020).
- [17] Waqas, H., Imran, M.: #CampFireMissing: An Analysis of Tweets About Missing and Found People from California Wildfires. *Proc. ISCRAM 2019* (2019).
- [18] Rust, P., Pfeiffer, J., Vulić, I., Ruder, S., Gurevych, I.: How Good is Your Tokenizer? On the Monolingual Performance of Multilingual Language Models. In: *Proc. 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, pp. 3118–3135. Association for Computational Linguistics, Online (2021).
- [19] Hu, J., Ruder, S., Siddhant, A., Neubig, G., Firat, O., Johnson, M.: XTREME: A Massively Multilingual Multi-task Benchmark for Evaluating Cross-lingual Generalisation. In: *Proc. 37th International Conference on Machine Learning*, pp. 4411–4421. PMLR (2020).
- [20] Yang, J., Liang, S., Zhang, Y.: Design Challenges and Misconceptions in Neural Sequence Labeling. In: *Proc. 27th International Conference on Computational Linguistics*, pp. 3879–3889. Association for Computational Linguistics, Santa Fe, New Mexico, USA (2018).
- [21] Paul, U., Ermakov, A., Nekrasov, M., Adarsh, V., Belding, E.: #Outage: Detecting Power and Communication Outages from Social Networks. In: *Proc. The Web Conference 2020*, pp. 1819–1829. ACM, Taipei Taiwan (2020).
- [22] Matsuda, K., Sasaki, A., Okazaki, N., Inui, K.: Annotating Geographical Entities on Microblog Text. In: *Proc. 9th Linguistic Annotation Workshop*, pp. 85–94. Association for Computational Linguistics, Denver, Colorado, USA (2015).
- [23] Utsu, K., Uchida, O.: Analysis of Rescue Request and Damage Report Tweets Posted During 2019 Typhoon Hagibis. *IEICE Transactions on Fundamentals*, E103-A, pp.1319-1323, (2020).
- [24] Devlin, J., Chang, M.-W., Lee, K., Toutanova, K.: BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. In: *Proc. 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, pp. 4171–4186. Association for Computational Linguistics, Minneapolis, Minnesota (2019).
- [25] Lan, Z., Chen, M., Goodman, S., Gimpel, K., Sharma, P., Soricut, R.: ALBERT: A Lite BERT for Self-supervised Learning of Language Representations. *Arxiv*. (2020).
- [26] Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., Levy, O., Lewis, M., Zettlemoyer, L., Stoyanov, V.: RoBERTa: A Robustly Optimized BERT Pretraining Approach. *ArXiv*. (2019).
- [27] Chalkidis, I., Fergadiotis, M., Malakasiotis, P., Aletras, N., Androutsopoulos, I.: LEGAL-BERT: The Muppets straight out of Law School. In: *Findings of the Association for Computational Linguistics: EMNLP 2020*, pp. 2898–2904. Association for Computational Linguistics, Online (2020).
- [28] Alsentzer, E., Murphy, J., Boag, W., Weng, W.-H., Jindi, D., Naumann, T., McDermott, M.: Publicly Available Clinical BERT Embeddings. In: *Proc. 2nd Clinical Natural Language Processing Workshop*, pp. 72–78. Association for Computational Linguistics, Minneapolis, Minnesota, USA (2019).
- [29] Wang, L., Gao, C., Wei, J., Ma, W., Liu, R., Vosoughi, S.: An Empirical Survey of Unsupervised Text Representation Methods on Twitter Data. In: *Proc. 6th Workshop on Noisy User-generated Text (W-NUT 2020)*, pp. 209–214. Association for Computational Linguistics, Online (2020).
- [30] Nguyen, D.Q., Vu, T., Tuan Nguyen, A.: BERTweet: A Pre-trained Language Model for English Tweets. In: *Proc. 2020 Conference on Empirical Methods in Natural Language Processing: System Demonstrations*, pp. 9–14. Association for Computational Linguistics, Online (2020).

## A ツイートの収集

災害種	災害名	開始時刻	収集したツイート数	絞り込み後
豪雨	平成 30 年 7 月豪雨	2018/07/06 11:10	1,729,800	28,019
	令和 2 年 7 月豪雨	2020/07/03 22:50	714,771	14,189
	令和 3 年 8 月豪雨	2021/08/13 02:45	549,342	12,458
台風	平成 30 年台風第 21 号	2018/09/04 06:00	4,608,594	41,113
	令和元年房総半島台風	2019/09/08 23:00	2,475,831	32,619
	令和元年東日本台風	2019/10/12 13:00	10,644,221	78,599
地震	平成 28 年熊本地震	2016/04/14 21:26	3,739,730	41,462
	平成 30 年大阪北部地震	2018/06/18 07:58	4,929,317	51,969
	平成 30 年北海道胆振東部地震	2018/09/06 03:07	4,868,427	59,621

## B ツイート収集時のクエリ

地震 OR 大雨 OR 火災 OR 国道 OR 津波 OR 突風 OR 強風 OR 暴風 OR 横転 OR 高潮 OR 高波 OR (運転 (取りやめ OR 取り止め OR 見合わせ OR 見合せ OR 見あわせ)) OR 運休 OR 通行止 OR 通行不 OR 通行規制 OR 渋滞 OR 避難 OR 救助 OR 通信障害 OR 停電 OR 帰宅困難 OR 帰宅難民 OR 足止め OR 帰れない OR 帰れなくなった OR 地滑り OR 山崩れ OR 土砂崩れ OR 倒木 OR 落石 OR 落木 OR 崖崩れ OR がけ崩れ OR 土砂流入 OR 地崩れ OR 冠水 OR 氾濫 OR 水没 OR 増水 OR 浸水 OR 洪水 OR 断水 OR 倒壊 OR 全壊 OR 被害状況 OR 崩壊 OR 崩落 OR 半壊

## C アノテーションした場所参照表現の例

種別	場所参照表現の例	ツイートの例 (著作権の観点から実在のツイートを加工)
地名	岩手県 倉敷市	大雨でスパークした電柱動画は倉敷市周辺で、9/4 14:20 頃に撮影
施設名	博多駅 東海大学	札幌地震による停電?札幌駅付近のマンションでは停電しました
路線名	京浜東北線 山手線	【東武亀戸線上下線 終日運転見合わせ】帰れねえ、、、
河川名	球磨川 阿武隈川	球磨川が氾濫寸前とのことですが、家の前の川はもう溢れそう #救助
道路名	国道 4 号線 東北道	国道 4 号線の一部で大規模な土砂崩れが発生し通行止め。この有様 <URL>
総称	病院 コンビニ	病院周辺の道路は水没しているため現在孤立状態にあります。

## D 同じ種類の災害で学習と推論を行った場合の実験結果

災害種	実験 ID	学習	推論	適合率	再現率	F 値
豪雨	(1)	令和 3 年 8 月豪雨	令和 3 年 8 月豪雨	0.740	0.782	0.761
	(2)	令和 2 年 7 月豪雨		0.696	0.689	0.693
	(3)	平成 30 年 7 月豪雨		0.658	0.658	0.658
	(4)	豪雨混合(2+3)		0.722	0.741	0.731
地震	(5)	平成 30 年北海道胆振東部地震	平成 30 年北海道胆振東部地震	0.768	0.809	0.788
	(6)	平成 28 年熊本地震		0.774	0.766	0.770
	(7)	平成 30 年大阪北部地震		0.755	0.739	0.747
	(8)	地震混合(6+7)		0.718	0.691	0.705
台風	(9)	令和元年東日本台風	令和元年東日本台風	0.734	0.793	0.762
	(10)	平成 30 年台風 21 号		0.678	0.778	0.725
	(11)	令和元年房総半島台風		0.628	0.692	0.659
	(12)	台風混合(10+11)		0.665	0.773	0.735

## E 異なる種類の災害で学習と推論を行った場合の実験結果

	Target(豪雨)			Target(台風)			Target(地震)		
	適合率	再現率	F 値	適合率	再現率	F 値	適合率	再現率	F 値
Target	0.740	0.782	0.761	0.734	0.793	0.762	0.768	0.809	0.788
豪雨	0.722	0.741	0.731	0.672	0.682	0.677	0.771	0.681	0.723
台風	0.696	0.699	0.698	0.665	0.773	0.715	0.735	0.739	0.737
地震	0.638	0.720	0.676	0.655	0.737	0.694	0.718	0.691	0.705