

救助要請抽出のための災害時ツイートの解析

山根 有輝也^{†1} 小泉 佑揮^{†1} 長谷川 亨^{†1}

概要: 2018年の西日本豪雨をはじめ、災害時に救助要請を表明するために SNS、とりわけ Twitter が利用されつつある。現実には、「救助要請」などのキーワードを含んでいながら救助要請と関連のないツイートも多い。本研究では救助要請の機械的な抽出に向けて、現実の災害時のツイート、具体的には 2018 年西日本豪雨災害におけるツイートを解析する。さらに、2019 年台風 19 号における救助要請関連ツイートを対象に、リカレントニューラルネットワークによる救助要請の抽出実験を実施し、機械学習による救助要請抽出の課題を検討する。

キーワード: ソーシャルメディア、Twitter、災害、分析、機械学習

Abstract: During catastrophic disasters like the Japan floods 2018, phone-based emergency call systems may not work as expected due to heavy congestion or network disruption. Social media, e.g., Twitter and Facebook, has been playing an important role as a communication tool to deliver rescue requests, which complements phone-based emergency call systems, in disasters. Understanding rescue requests on social media is key to realizing the automatic extraction of rescue requests from a vast amount of social media posts. This paper analyzes rescue-related tweets, which are tweets containing rescue-related keywords like #rescue, and proposes a taxonomy for rescue-related tweets. According to the analysis, we construct a classifier based on a recurrent neural network and an annotation mechanism to identify why the classifier identifies a tweet as a rescue request or not. Using these two models, we reveal that the reason why there are many false negatives, i.e., the number of rescue requests identified as non-rescue requests.

Keywords: Social Media, Twitter, Disaster, Analysis, Machine Learning

1. はじめに

災害時には、救助が必要な人からの救助要請を適切な部署あるいは初動対応員に送信することが、効率的な災害対応の鍵である。しかし、大規模災害時には、ネットワークインフラストラクチャーの障害や輻輳により 119 番などの緊急通信が利用できないことがある [1]。日本においても、例えば西日本豪雨災害の際、倉敷市消防局の電話台への通報が通常時の平均的な通報件数の 30 倍以上にのぼったため、緊急通信サービスが輻輳し、利用できなかったことが報告されている [2]。

その一方で、大規模災害時の救助要請のあり方が変化しつつある。2018 年の西日本豪雨災害 [3] や 2019 年の台風 19 号 [4] では、多くの被災者が SNS、とりわけ Twitter を用いて救助要請を表明した。2 章で議論するように、本研究においても西日本豪雨災害の際には、Twitter 上で 312 地点の救助要請が観測された。しかし、これらの救助要請はただちに実際の救助活動へ活用できるわけではない。西日

本豪雨災害の際には、多くの救助要請が Twitter 上にポストされたものの、そのうち実際の救助活動につながったツイートは存在しなかったことが示されている [2]。一方で、台風 19 号の際には、長野県において手動で収集された救助要請ツイートが、約 50 件の救助活動に活用された事例が報告されている [5]。しかし、現状では多くの人員をツイートの分析に割く必要があり、緊急時には救助要請ツイートを機械的に抽出する技術が必要である。

ハリケーンなどの大規模災害に関連するツイートの分析については、すでにいくつかの研究が存在する [6-8]。しかし、これらの研究では救助要請のツイートについては分析されていない。これに対して、本稿では、救助要請関連ツイートを分析する。

本稿の貢献は以下の通りである。まず、救助要請関連ツイートを分析し、様々なコンテキストのツイートが含まれることを明らかにするとともに、それを救助要請、災害情報、同情、災害への便乗などのカテゴリに手動で分類した。次に、ツイートはフリーフォームテキストである特性上、個々の表現レベルの格差や表現のあいまい性を伴っており、それらが機械学習を用いた分類を困難にしていることを示

^{†1} 現在、大阪大学 大学院情報科学研究科
Presently with Osaka University

す。最後に、機械学習を用いた分類実験とアノテーションを用いた分析により、機械学習による救助関連ツイートの分類が困難である理由を調査した。

本稿の構成は以下の通りである。2章で救助要請関連ツイートを含む災害関連ツイートを分析する。3章で、ニューラルネットワークを用いて救助要請関連ツイートを分類する実験を行うとともに、アノテーションを用いて、救助要請関連ツイートの表現のあいまい性により生じる課題や、機械学習による分類が失敗した理由について考察する。4章で関連研究を紹介し、最後に5章で本稿をまとめる。

2. 災害関連ツイートの分析

本章では、災害関連ツイートおよび救助要請関連ツイートを分析する。ここで、災害関連ツイートと救助要請関連ツイートは、それぞれ災害に関連するキーワード、救助要請に関連するキーワードを含むツイートと定義する。それぞれのキーワードについては後述する。

2.1 災害関連ツイートデータセット

本章の解析のため、2018年に発生した西日本豪雨災害[3]のツイートを収集した。収集には、Twitter search API [9]を用い、表1に示すキーワードを指定してツイートを収集した。話題別のツイート数を解析するため、表1に示す5つのトピック、つまり、豪雨災害、救助要請、救助隊、インフラストラクチャー、ボランティアに関するキーワードを指定した。ただし、これらのツイートの内容とトピックが一致するとは限らない。例えば、救助要請に関するキーワードを含んでいながら救助要請ではないツイートも存在する。したがって、これらのツイートを関連ツイートと呼称する。具体的には、それぞれのトピックのキーワードを1つでも含むツイートを、豪雨災害関連ツイート、救助要請関連ツイート、救助隊関連ツイート、インフラストラクチャー関連ツイート、ボランティア関連ツイートと呼ぶ。また、これらのツイートをまとめて災害関連ツイートと呼ぶ。

上記の方法で収集した災害関連ツイートの総数は、6,978,389である。Twitter APIを用いたリツイートは除外している。各トピックのツイート数を表2にまとめる。

2.2 災害関連ツイートの時系列分析

はじめに、西日本豪雨災害での洪水による被害状況と災害関連ツイートとの関連を理解するため、1日あたりのツイート数の推移を整理した。

結果を図1aに示す。横軸は日付であり、縦軸は1日あたりのツイート数である。結果を理解するために、西日本豪雨災害の被害状況を簡単に述べる。2018年6月末頃から7月中旬にかけて、西日本を中心に激しい雨が降り続いた結果、西日本の多くの地域で河川の氾濫が相次いだ。さらに、7月5日から7日にかけて、岡山県、広島県や愛媛県

表1 災害関連ツイート収集に用いたキーワード

トピック	キーワード
豪雨災害	豪雨、豪雨災害、台風、西日本豪雨、災害、水害、浸水、被災、被災地、洪水、決壊、避難、安否、安否確認
救助要請	救助、救援、助けて、たすけて、SOS、レスキュー
救助隊	レスキュー、自衛隊、消防、病院、警察、自治体
インフラストラクチャー	インフラ、ライフライン、水道、電気、ガラクチャー、障害、復旧、繋がらない、繋がりにくい、119、110
ボランティア	ボランティア、救援物資、支援物資

表2 各トピックのツイート数

トピック	ツイート数
災害関連ツイート	6,978,389
救助要請関連ツイート	246,807
救助隊関連ツイート	932,605
インフラストラクチャー関連ツイート	889,889
ボランティア関連ツイート	324,935

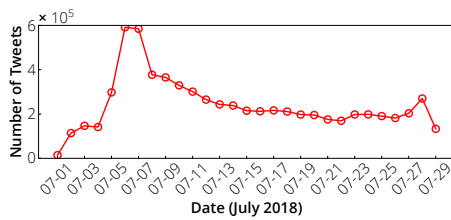
を中心に激しい洪水が発生し、西日本豪雨災害における被害をより深刻なものにした。

図1aに、災害関連ツイートの時系列変化を示す。岡山県や広島県における洪水の激化と一致するように、7月6日から7日にかけて災害関連ツイートの数が急激に増加し、それぞれ591,514件と584,692件にのぼった。次に、図1bに救助要請関連、救助隊関連、インフラストラクチャー関連、ボランティア関連のツイート数の推移を示す。豪雨災害関連のツイートは災害関連ツイートとの差が小さいため省略した。救助要請関連ツイートの数は、洪水の激化に伴い急激に上昇したことが確認できる。一方で、救助隊関連やインフラストラクチャー関連のツイートは災害発生中、災害後を問わず定常的に発生していた。そしてボランティア関連のツイートは災害後から増加し、復旧期間を通じて定常的に発生していることを確認した。

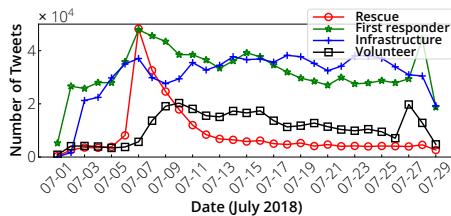
2.3 救助要請の発生地点の分析

次に、実際の救助要請の確認とその発生地点について分析した。ツイートには住所や経緯度などのメタデータを付与することも可能であるが、西日本豪雨災害時には詳細な地点を示すメタデータが付与されている救助要請が存在しなかった。そのため、具体的な住所や建物名などを一意に特定できる情報と結びつく救助要請を手動で抽出した。リツイートに加え、コピーされた情報や伝聞情報などを排除すると、312地点で救助要請が発生していることが分かった。図2に、救助要請の発生地点をプロットした。プライバシーに配慮して、縮尺の小さい地図を掲載した。

本解析から分かる重要なポイントの1つは、洪水の発生



(a) 災害関連ツイート総数



(b) 救助要請、救助隊、インフラストラクチャー、ボランティア関連ツイート数

図 1 1日あたりのツイート数の推移

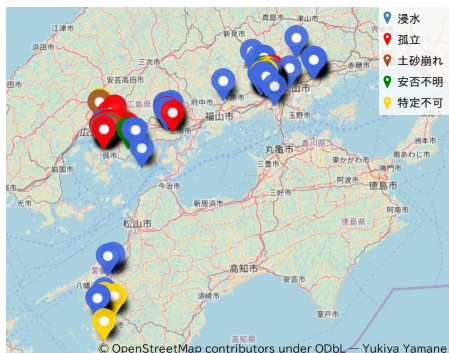


図 2 西日本豪雨時の救助要請に示されていた地点

した地域を中心に、限られた領域で救助要請が発生していることである。ある地域では、7.5 km × 5 km の範囲の 235 地点において救助要請が発生した。このように局所的に発生した多数の救助要請こそ、119 番通報の輻輳 [2] の要因の 1 つと推測できる。もう 1 つの重要なポイントは、救助要請関連ツイート数に対して具体的な住所と結びついた救助要請の数が 312 と非常に少ないという点である。これについて分析するため、次節で救助要請関連ツイートの内容に基づいた分類を行う。

2.4 救助要請関連ツイートの分類

救助要請関連ツイート数に対して救助要請数が少ないという分析結果を踏まえ、この理由を明らかにするために救助要請関連ツイートを手動で分類した。救助要請関連ツイート全数をチェックすることは難しいため、救助要請関連のハッシュタグ (#救助、#救助要請、#救援、#救援要請、#助けて、#助けてください、#助けて下さい) を含むツイート 5,304 件を抽出し、それらを読んだ結果、以下の 8 カテゴリーに分類できた。

- 救助要請：救助要請であり、被災状況と被災地点を含

むツイート

- 不完全な救助要請：救助要請であるものの、被災状況あるいは被災地点などの情報が欠損しているツイート
- 災害情報：災害の状況を伝えるツイート
- ボランティア：ボランティア活動、寄付や支援物資の提供あるいは要求を呼びかけるツイート
- アドバイスと補助情報：被災者に対するアドバイスや既存ツイートに対して情報を付与するツイート
- 同情：被災者に対する同情や祈りのツイート
- 便乗：災害に関連のない情報をツイートするために意図的に救助要請関連キーワードを利用したツイート
- 無関係：災害に関連のないツイート

この分類は、救助要請に関連するキーワードでツイートを抽出したとしても救助要請に関連のないツイートが多く含まれるため、キーワード検索による救助要請の抽出が難しいことを示唆している。

西日本豪雨災害の最中の 7 月 10 日に、Twitter 公式の Twitter ライフラインアカウント (@TwitterLifeline) で、Twitter 上で救助要請を発するときの勧告が出された [10]。これによると、救助要請をポストする際には、#救助要請の付与に加え、具体的な被災状況と被災地点あるいは GPS 情報の記載と、可能な場合には被災状況の写真を添付することが推奨されている。本研究では写真データを対象としていないため、上記の勧告から写真の有無を除いた項目に基づいて、被災状況と地点の両方を含む救助要請と、その 2 つのいずれかが欠損した救助要請に分類した。

最後に、便乗と無関係のカテゴリーについては、いずれも災害と関連のない点で共通しているが、便乗のカテゴリーには広告や政治的主張などを広めるために救助要請関連のハッシュタグを悪用したツイートが含まれ、無関係のカテゴリーには、ゲームや映画などのストーリー上の救助など、災害との関連がないものの、意図的な文脈の操作もないツイートが含まれる。本研究の対象外ではあるが、機械的にツイートを分類する際にはこれらの表現の差が課題となる可能性がある。

3. 機械学習を用いた分類実験

3.1 データセット

2.1 章のデータセットに加え、同様の方法で 2019 年の台風 15 号 [11]、2019 年の台風 19 号 [4] の発生前後にポストされた約 2,300 万件の災害関連ツイートを収集した。その中から、次の 7 種のハッシュタグ (#救助、#救助要請、#救援、#救援要請、#助けて、#助けてください、#助けて下さい) を含む救助要請関連ツイートを抽出した。さらに、台風 19 号では被害の大きかった長野県において利用された「#台風 19 号長野県被害」も含めた。西日本豪雨災害のデータは 5,304 件、台風 19 号時のデータは 2,091 件であった。

このうち、西日本豪雨災害時のツイートを教師データ、

台風 19 号時のツイート进行测试データとして用いた。

3.2 分類器の構成

3.2.1 ツイートへのラベル付与

2 章の分類に基づいて、ツイート t に対してラベル y_t を付与する。比較対象として二値ラベル $z_t = \{0, 1\}$ も付与する。1 は救助要請、0 は救助要請でないことを表す。それぞれのラベルを用いて生成する分類器を、以降それぞれ 8 種分類器、二値分類器と呼ぶ。

3.2.2 ツイートの事前処理

MeCab [12] を利用し、ツイート本文を形態素解析し単語に分割する。このとき、活用語尾のある単語は終止形に揃える。次に、URL、絵文字、記号を除外する。ただし、スクリーンネームは、他者とのやり取りなどの有無を読み取るうえで重要であるが、個々のスクリーンネームは重要ではないため、*SCREENNAME* という特別な単語に置換する。以上により、ツイート t を単語 $\{w_{1,t}, \dots, w_{n,t}\}$ に分割する。また、分析対象とするツイートの最大長は 100 単語とした。

3.2.3 単語の分散表現の生成

単語の分散表現の生成には、GloVe [13] を用いた。GloVe は単語の共起関係に基づいた単語のベクトル化を可能とする。分散表現を生成するためのコーパスには、災害関連ツイート約 2,300 万件を利用し、それに 2 回以上含まれる 651,423 単語をベクトル変換の対象とした。以上により、単語 $w_{i,t}$ がベクトル $x_{i,t} \in \mathbb{R}^{256}$ に変換される。

3.2.4 ニューラルネットワークによる分類

リカレントニューラルネットワーク (RNN) を用いて分類器を生成する。分類器の構成の概要を図 3 に示す。RNN はニューラルネットワークの一種であり、文脈などの時系列データの解析に特化している。本研究では RNN の一種である長・短期記憶 (LSTM: Long short-term memory) [14] を用いる。さらに、Kshirsagar ら [15] による Twitter の本文解析と同様、ツイートを構成する短いテキストを処理するため、双方向 LSTM を用いた。ツイート内の各単語を順方向に入力して学習させる第一の LSTM 層だけでなく、逆方向に入力した第二の LSTM 層を併用し、2 層の結果を結合する。

3.2.5 ニューラルネットワークによるアノテーション

RNN による分類結果の根拠を明らかにするため、Kshirsagar ら [15] と同様にアノテーション手法を使い、二値分類器の分類過程で導出される隠れ状態を分析する。図 3 において、ツイート t の各単語を入力したときの LSTM の出力 $a_{i,t}$ が隠れ状態の値であり、以降アノテーションと呼ぶ。RNN による判定の理由を分析するため、各単語を入力したときのアノテーション値により RNN がどの表現で救助要請と判定しているのかを分析する。隠れ状態の値が高いほど、その表現が救助要請と関連が強いことを意味する。

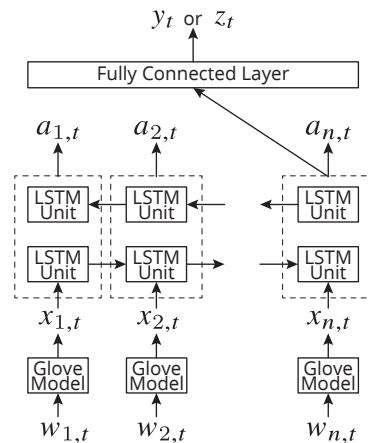


図 3 分類器の構成

表 3 二値分類器の結果

予測 \ ラベル	1	0
1	181	70
0	100	1740

表 4 8 種分類器の結果

予測 \ ラベル	救助要請	救助要請でない
救助要請	130	21
不完全な救助要請	69	81
災害情報	40	199
同情	21	934
アドバイスと補助情報	17	323
ボランティア	1	24
災害への便乗	2	139
災害と無関係	1	89

救助要請の分類器を生成する際には、LSTM の出力を全結合層に通したあと、ラベル y_t あるいは z_t を用いて学習させた。一方、アノテーションの分類器の生成では、LSTM の各出力層に対して救助要請か否かを示すラベル z_t を用いて学習させた。

3.3 ニューラルネットワークによる分類結果

はじめに、二値分類器による分類結果を表 3 に示す。正解率は 0.919 であるが、再現率は 0.644、適合率は 0.721 であり、F 値は 0.680 であった。このうち本稿で重視するのは再現率、すなわち、救助要請のうち正しく救助要請と判定されたツイートの割合であり、再現率が高いほど救助要請の見落としが少ないことを意味する。裏を返せば、この再現率は、二値分類器において約 36% もの救助要請が見落とされることを示している。

次に 8 種分類器による分類結果を表 4 に示す。2.4 章で定義したカテゴリーに該当すると判定されたツイートの件数をまとめるとともに、それぞれについて、実際の救助要請とそうでないツイートに分類した。再現率は 0.708 であり、救助要請の見落としは二値分類器の場合よりも少ないものの、依然として計 82 件にのぼる。特に、救助要請であ

表5 救助要請の判定に寄与した単語

分類	単語
被災状況	屋根の上、取り残された、孤立、陸の孤島、浸水
被災対象	父、母、祖父母、家族、年寄り、子供

るものの、災害情報、同情、アドバイスと補助情報のカテゴリへ分類される件数が多い。次節では、アノテーションを用いて救助要請の見落としの要因を分析する。

3.4 アノテーションによる誤判定理由の分析

救助要請と判定されたツイートについてアノテーション値 ($a_{i,t}$) を算出し、値が高い表現中に頻繁に出現した単語をピックアップしたものを表5にまとめた。これは教師データとして用いた西日本豪雨災害時の岡山県における被災状況と一致している。

救助要請の分類で誤判定となる要因はいくつかある。やはり、教師データとして用いた西日本豪雨災害時とテストデータとして用いた台風19号時の被災状況の違いから、西日本豪雨災害時には存在しなかった表現が含まれる救助要請が、非救助要請と判定される例が多かった。一方、アノテーションを利用して判明した要因の1つに、救助要請の中に、救助要請と関連のない文脈の表現が混入している例がみられた。例えば、「疲れているのか母と弟はお昼寝中。救助要請してからかなりたちました。(中略)119番も『見付けてもらえるまでがんばれ』としか言いませんでした。XXX町には、まだ何人も残されてます。#台風19号長野県被害」というツイートでは、最初の一文が救助要請に関連のないコンテキストであり、この周囲のアノテーション値が低いことが確認できた。これにより、ツイート全体が非救助要請と判定された可能性が高い。

4. 関連研究

本章では、本研究の関連研究を災害時のSNS上の情報解析、災害時のSNS情報の利用、および機械学習を用いたSNS上の情報の処理技術に分類してまとめる。

Twitterを対象とした災害に関連するSNSポストの分析については、多くの研究がなされている。Alamら[7]は、2017年にアメリカで発生した3つのハリケーン(Harvey、Irma、Maria)発生時にポストされた災害関連ツイートを解析している。具体的には、災害関連ツイートのテキストとマルチメディアデータの分析を実施している。テキスト解析の結果、災害関連ツイートをいくつかのカテゴリに分類している。ツイートの分類方法については、Alamら[7]の方法に基づいている。Yangら[8]もハリケーンHarvey発生時にポストされたツイートを分析し、そこから抽出できるイベントレポートの信憑性の推定フレームワークを検討している。いずれも、災害関連のツイートを広く扱った研究であるのに対して、本研究では救助要請関連ツイートを

対象に分析するものである。

災害対応について、SNS上の情報の利用が検討されている。我々は、これまでの研究[16]において、災害時の救急隊などの初動対応員やボランティアに、SNS上の信憑性の高い情報を提供するためのフレームワークを提案している。このなかで、SNS上の情報を適切な初動対応員に転送することを目的として、機械学習を用いてSNS上の情報をイベントに分類する方法を提案している。本研究の動機は文献[16]と同じであるが、文献[16]では災害情報の運用に焦点を当てているのに対して、本研究では救助要請の分類に焦点を当てている。

最後に、機械学習を用いたSNS上の情報の処理技術をまとめる。ニューラルネットワークを用いたSNS上のフェイク情報の判定技術に関する研究がなされている[17,18]。これらの研究はテキスト情報に加え、同じ事象に関するSNSポストのインターバルなどの特徴量を用いることでフェイク情報の抽出を目指すものである。一方、Kshirsagar[15]は、SNSポストのテキスト情報を解析し、自傷、自殺、摂食障害などの危機に繋がるSNSポストの判定技術と、それらの危機に直結する表現を抽出する技術を提案している。本研究のツイート分類については、Kshirsagarらのモデルをベースにしている。

5. おわりに

大規模災害時の救助要請表明にSNS、とりわけTwitterが利用されつつある。しかし、救助要請に関するキーワードを含んでいながら救助要請ではないツイートも多く、Twitter上の救助要請の活用にあたり、機械的な抽出には課題が残る。そこで本研究では、西日本豪雨災害時の救助要請関連ツイートを解析し、救助要請関連のキーワードを含んでいながら救助要請ではないツイートが多く存在することを明らかにするとともに、それらを8つのカテゴリに分類した。リカレントニューラルネットワークによる分類実験では、8種の分類に基づく抽出の方が救助要請か否かの二値分類よりも再現率が高いことを示した。最後に、リカレントニューラルネットワークのアノテーション手法を用いて、誤った分類となる理由を分析した。

謝辞

本研究はNICT受託研究課題193によるものである。

参考文献

- [1] Jahanian, M., Xing, Y., Chen, J., Ramakrishnan, K. K., Seferoglu, H. and Yuksel, M.: The Evolving Nature of Disaster Management in the Internet and Social Media Era, *Proceedings of IEEE International Symposium on Local and Metropolitan Area Networks* (2018).
- [2] 日本放送協会(NHK): 災害時の「#救助要請」は救助につながったのか, <http://bit.ly/2vJiQ7Z> (2018).

- [3] Wikipedia: 平成 30 年 7 月豪雨, <http://ja.wikipedia.org/w/index.php?curid=3810926>.
- [4] Wikipedia: 令和元年台風第 19 号, <http://ja.wikipedia.org/w/index.php?curid=4015940>.
- [5] 日本放送協会 (NHK): 長野県ツイッターの救助要請収集で約 50 件救助に, <https://bit.ly/3741noG> (2019).
- [6] Takahashi, B., Tandoc, E. and Carmichael, C.: Communicating on Twitter during a disaster: An analysis of tweets during Typhoon Haiyan in the Philippines, *Computers in Human Behavior*, Vol. 50, pp. 392–398 (2015).
- [7] Alam, F., Ofli, F., Imran, M. and Aupetit, M.: A Twitter Tale of Three Hurricanes: Harvey, Irma, and Maria, *Proceedings of International Conference on Information Systems for Crisis Response and Management* (2018).
- [8] Yang, J., Yu, M., Qin, H., Lu, M. and Yang, C.: A Twitter Data Credibility Framework—Hurricane Harvey as a Use Case, *International Journal of Geo-Information*, Vol. 8, No. 3 (2019).
- [9] Twitter, Inc.: Twitter API Documentation—Search Tweets, <https://developer.twitter.com/en/docs/tweets/search/api-reference/get-search-tweets>.
- [10] Twitter, Inc. (@TwitterLifeline): Twitter post, <https://twitter.com/TwitterLifeline/status/1016519147738419201> (2018).
- [11] Wikipedia: 令和元年台風第 15 号, <http://ja.wikipedia.org/w/index.php?curid=4000816>.
- [12] MeCab: Yet Another Part-of-Speech and Morphological Analyzer, <https://taku910.github.io/mecab/>.
- [13] Pennington, J., Socher, R. and Manning, C. D.: GloVe: Global Vectors for Word Representation, *Proceedings of Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pp. 1532–1543 (2014).
- [14] Hochreiter, S. and Schmidhuber, J.: Long Short-Term Memory, *Neural Computation*, Vol. 9, pp. 1735–1780 (1997).
- [15] Kshirsagar, R., Morris, R. and Bowman, S. R.: Detecting and Explaining Crisis, *Proceedings of ACM Workshop on Computer Linguistics and Clinical Psychology*, pp. 66–73 (2017).
- [16] Jahanian, M., Hasegawa, T., Kawabe, Y., Koizumi, Y., Magdy, A., Nishigaki, M., Ohki, T. and Ramakrishnan, K. K.: Di-ReCT: Disaster Response Coordination with Trusted Volunteers, *Proceedings of International Conference on Information and Communication Technologies for Disaster Management* (2019).
- [17] Ma, J., Gao, W., Mitra, P., Kwon, S., Jansen, B. J., Wong, K.-F. and Cha, M.: Detecting rumors from microblogs with recurrent neural networks, *Processings of International Joint Conference on Artificial Intelligence*, pp. 3818–3824 (2016).
- [18] Ruchansky, N., Seo, S. and Liu, Y.: CSI: A hybrid deep model for fake news detection, *Processings of International Conference on Information and Knowledge Management*, pp. 797–806 (2017).