

# 深層学習を用いた大規模災害の行動促進情報の分析

米田 吉希<sup>†</sup>   鈴木 優<sup>††</sup>   灘本 明代<sup>†††</sup>

<sup>†</sup> 甲南大学大学院自然科学研究科 〒 658-8501 兵庫県神戸市東灘区岡本 8-9-1

<sup>††</sup> 岐阜大学工学部 〒 501-1193 岐阜県岐阜市柳戸 1-1

<sup>†††</sup> 甲南大学知能情報学部知能情報学科 〒 658-8501 兵庫県神戸市東灘区岡本 8-9-1

E-mail: <sup>†</sup>m1824003@s.konan-u.ac.jp, <sup>††</sup>ysuzuki@gifu-u.ac.jp, <sup>†††</sup>nadamoto@konan-u.ac.jp

あらまし 近年, Twitter 上で災害に関する様々な情報交換が盛んに行われている. その情報の中でも, 行動促進ツイートと呼ばれる注意喚起やアドバイスのような行動を促す情報は, 人々の行動に影響を与える. また, 行動促進ツイートに誤った情報が含まれている場合, 閲覧ユーザが不利益を被る可能性がある. そこで, 大規模災害時の行動促進ツイートを抽出し, ユーザへの行動促進ツイートに対する警告の提示を行う必要があると考えた. 本論文では, 災害時の行動促進ツイートを抽出する手法として, LSTM, BLSTM, BERT の 3 つの深層学習の手法を比較検討する. さらに, 2019 年台風 15 号の大規模台風時と 2018 年の台風時の行動促進ツイートの比較分析を行う.

キーワード Twitter, 災害, 情報抽出, 行動促進, 深層学習, LSTM, BLSTM, BERT

## 1 はじめに

近年, 日本では多くの災害が発生している. このような大規模災害時には Twitter 上で様々な情報交換が盛んに行われている. その情報の中でも, 注意喚起やアドバイスのような行動を促す情報は, 人々の行動に影響を与えることから重要であると考えられる. 例えば, 「外から飛んでくる物に備え, カーテンやブラインドを閉めておきましょう」や「無理をして手作業で片付けようとするのはやめましょう」といったツイートが挙げられる. 本研究では, このような閲覧ユーザに行動を促す情報を行動促進情報と呼び, その行動促進情報を含むツイートを行動促進ツイートと呼ぶ[1].

行動促進ツイートは大きく分けて, 明示的行動促進ツイートと暗示的行動促進ツイートの 2 種類に分類される. 明示的行動促進ツイートとは, 例えば「台風が接近してるので河川には近寄らないようにしてください」のようにツイート内に明らかに行動を促進している文言を含むツイートである. 一方, 暗示的行動促進ツイートとは, ツイート内に明らかに行動を促進している文言は存在しないが, 他人に行動を促しているツイートである. 例えば「台風 15 号の接近により河川の増水等の恐れがあります」というツイート内には, 行動を促している文言は存在していないが, 書かれていることから想像できる行動をさせようとするように, 暗示的に行動を促している. 本研究において, 我々が対象とする行動促進ツイートは, 他人に行動を促進する内容を含むツイートのうち, 明示的行動促進ツイートとする.

また, 行動促進ツイートは, 災害時に行動促進ツイートを閲覧したユーザが実際にそのツイートの内容を行動に移すことで, 防災や減災, 被災地支援に繋がる可能性があるため重要であると考えられる. しかしながら, 大規模災害時は冷静な判断が困難であるため, 行動促進ツイートに誤った情報が含まれている場合, 閲覧ユーザが不利益を被る可能性がある. 東日本大震災

時のツイートを分析した研究[2]によると, 7,177 件のデマツイートのうち 53%の 3,773 件のツイートが行動促進ツイートであることがわかっている. この結果から, 災害時の行動促進ツイート内にデマ情報を含む可能性が高いことがわかる. そこで我々は, 大規模災害時の行動促進ツイートを抽出し, ユーザへの行動促進ツイートに対する警告の提示を行う必要があると考えた. これにより, ユーザが大規模災害時の正確な判断ができない状況において, 誤った情報を信じてしまうことにより, 不利益を被る事を減少できると考えられる.

本研究の手順は以下の通りである.

- (1) 大規模災害時の行動促進ツイートの抽出.
- (2) 抽出した行動促進ツイートに対しての真偽を判定.
- (3) 行動促進ツイートに対して警告の提示.

本論文では, その内 (1) の Twitter 上の大規模災害時の行動促進ツイートの抽出手法を提案し, 比較分析を行う.

これまで我々は, 災害時と平常時の行動促進ツイート抽出手法を提案し, 行動促進ツイートの比較分析を行ってきた [1] [3]. そして, 災害時と平常時の行動促進ツイート抽出には Long short-term memory(LSTM) が最も適していることを示している. しかしながら, 2019 年の台風 15 号のような大規模災害では, これまでにないほどの甚大な被害を受けたため, 投稿された行動促進ツイートはこれまでの台風時のものと異なると考えられる. そこで本論文では, 大規模災害として 2019 年の台風 15 号を対象とし, 災害時と平常時の行動促進ツイート抽出に最適な LSTM に加えて, 新たに Bidirectional LSTM(BLSTM) と Bidirectional Encoder Representations from Transformers(BERT) の 2 つの手法を提案する. 具体的には, それぞれの手法のハイパーパラメータに関して決定すると共に, 大規模災害時の行動促進ツイート抽出手法の比較実験を行い考察する. そして, 大規模災害時の行動促進ツイートの特徴分析を行う. さらに, 大規模災害とその他の災害との比較を行う事により, 我々の提案する手法が大規模災害にも有用であるかを検証

表 1 行動促進ツイートの例

【東京都#大島支庁から】各港への立入禁止及び車両移動についてお知らせします。台風 15 号の影響により各港の岸壁は波浪被害を受ける恐れがあります。大変危険ですので、近づかないようにしてください。また、各港の駐車場に駐車中の車両は移…
【通電火災に注意】台風被害による家屋や建物の損傷により、通電時に通電火災が起こる危険性があります。電力復旧前にすべての家電製品及びコンセントやケーブル等の破損状況を確認してください。(午後 4 時 00 分頃、防災行政無線にて放送)…
台風被害の復旧作業をしている皆さんへ。工務店の人や造園屋さんが来られないからといって無理をして手作業で片付けようとするのはやめましょう。倒木や倒壊した構造物の撤去は重機が入ればすぐ片付きます。作業中は熱中症にならないように気をつけてください。暑い時間帯は動かないで休みましょう。

する。具体的には、2019 年の台風 15 号と 2018 年の台風時の行動促進ツイートの比較分析を行う。

以下、2 章で関連研究について述べる。そして、3 章で 3 つの行動促進ツイート抽出手法について述べ、4 章で 3 つの提案手法を用いた比較実験を行い、その考察について述べる。最後に 5 章でまとめと今後の課題について述べる。

## 2 関連研究

Twitter から災害時の有益な情報を抽出する研究は数多くされている。Ashktorab ら [4] は、災害時のツイートから実用的な情報をリアルタイムに抽出するための手法を提案している。真中ら [5] は、災害に関するニュース投稿のリツイート直後の投稿に着目した災害関連情報の抽出手法を提案している。そして、ニュース投稿の特徴語句を用いた手法で高い適合率を示している。Rudra ら [6] は、Twitter 上の災害状況の情報を要約する手法を提案している。川崎ら [7] は、信頼性の高い救助要請の特徴を明らかにし、信頼性の高い救助要請ツイートを自動抽出するために、機械学習と住所情報をを用いた抽出手法を提案している。Cameron ら [8] は、Twitter 上の危機管理のための情報を検出、評価、要約、報告を行うシステムを構築している。湯沢ら [9] は、感動詞や代表語の共起関係を用いて、災害に関連する投稿を検索する検索語集合を自動的に抽出する手法を提案している。そして、異なる種類の大規模災害時のツイートに対して、提案手法による検索語集合の抽出が高い精度を示している。これらの研究は、Twitter 上から災害時の重要な情報を抽出するという点では本研究と類似しているが、本研究では大規模災害時の情報の中でも重要であると考えられる、他人に行動を促進するツイートに着目し研究を行っているという点で異なる。

さらに、Twitter 上の災害情報を分析する研究も多く行われている。北田ら [10] は、東日本大震災前後のツイートから得られた議論話題のトピックをフィルタリングとランキングを用いて可視化し、分析を行っている。その結果、Twitter における議論話題の変遷が観測でき、トピックの時系列関係を幅広く観察できることを示している。佐藤ら [11] は、2018 年西日本豪雨災害を対象に「#救助」ツイートに対する内容分析と過去の豪雨災害との比較分析を行っている。そして、救助要請を実際に求めている「#救助」によるツイートがごくわずかで、被災地外の不急の発信が依然として多いことを明らかにしている。榎田ら [12] は、災害発生直後のツイートやユーザの行動だけ

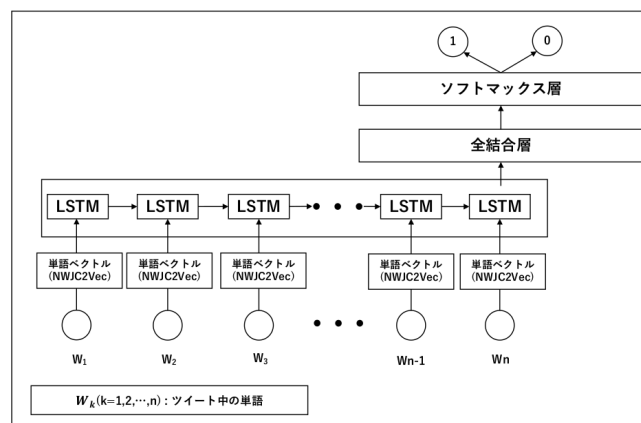


図 1 LSTM のネットワークの構造

でなく、日常的な防災・減災に着目し、災害情報を発信している認証済みアカウントのツイートとフォロワーの属性の分析を行っている。本研究では、他人に行動を促すツイートに着目し、大規模災害時における行動促進ツイートの特徴分析と大規模災害の 2019 年の台風 15 号時と 2018 年の台風時との比較分析を行う。

## 3 行動促進ツイートの抽出

本論文では大規模災害として 2019 年の台風 15 号を対象とし、LSTM, BLSTM, BERT を用いて災害時の行動促進ツイートを抽出する手法の検討を行う。大規模災害として 2019 年の台風 15 号時のツイートを対象とし、行動促進ツイートを 3 つの手法を用いて抽出し検討する。我々が抽出する大規模災害時の行動促進ツイートの例を表 1 に示す。また、それぞれの抽出手法の前処理として使用するデータ内から、url, Twitter のユーザ名を取り除く処理を行う。

### 3.1 LSTM を用いた抽出手法

単語の順序を考慮することのできる方法の一つである LSTM [13] を用いて行動促進ツイートの抽出を行う。LSTM の実装には機械学習ライブラリである Chainer<sup>1</sup> を使用する。使用する LSTM のネットワークの構造を図 1 に示す。特徴ベクトルは、ツイートに出現する全単語を fastText<sup>2</sup> を用いて分散表現を行ったものを使用する。fastText の学習済みモデルデー

1 : Chainer <https://chainer.org/>

2 : <https://fasttext.cc/>

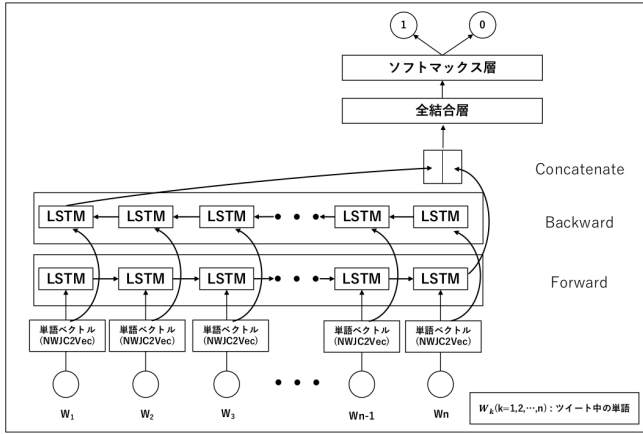


図2 BLSTMのネットワークの構造

タとして、国立国語研究所の NWJC2Vec [14] を利用する。使用する単語は、ツイートに対して形態素解析を行い、得られた全単語を用いる。全単語を用いる理由としては、文の単語の流れを考慮することにより、同様の表現でも違う意味が存在するという問題に対応することができると考えたためである。また、使用したモデルに存在していない未知語に対しては、今回は0ベクトルを使用した。

LSTMの各種ハイパーパラメータは、隠れ層は1、ベクトルサイズは300、バッチサイズは500、エポック数は20、学習率は0.001、ドロップアウト率は0.5、オプティマイザーはAdam [15] を使用する。LSTMモデルは1層で構成され、ユニット数はツイートに含まれる単語数である。300次元の特徴ベクトルはLSTMのユニットへの入力であり、LSTMのユニットの出力は次のユニットの入力となっている。最後のLSTMのユニットの出力は全結合層への入力となっている。この全結合層は2つの値を出力し、その値を用いてSoftMax層にて、行動促進ツイートか否かを判別する。

### 3.2 BLSTMを用いた抽出手法

3.1節で我々が提案したLSTMが考慮しているのは順方向のみの単語の順序である。そこで、順方向からの単語の流れだけでなく、逆方向からの情報も捉えることが出来るBLSTM [16] を用いて行動促進ツイートの抽出を行う。逆方向での出力を用意することで、後方の情報も考慮することができ、抽出精度が向上すると考えられる。BLSTMの実装にはLSTMと同様でChainerを使用する。使用するBLSTMのネットワークの構造を図2に示す。特徴ベクトルは、ツイートに出現する全単語をfastTextを用いて分散表現を行ったものを使用する。fastTextの学習済みモデルデータはLSTMで使用するモデルと同じモデルを用いる。使用する単語は、ツイートに対して形態素解析を行い、得られた全単語を用いる。

BLSTMの各種ハイパーパラメータは、隠れ層は2、ベクトルサイズは300、バッチサイズは500、エポック数は20、学習率は0.001、ドロップアウト率は0.5、オプティマイザーはAdamを使用する。入力ユニット数はツイートに含まれる単語数であり、BLSTMは順方向と逆方向それぞれ1層のLSTMモデル

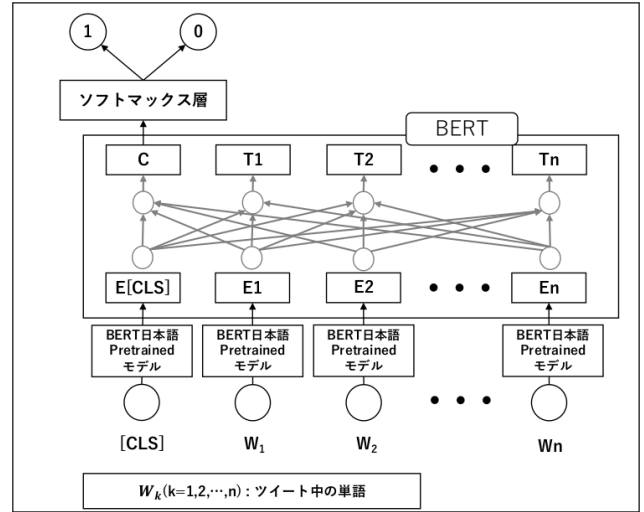


図3 BERTのネットワークの構造

で構成される。そして、それぞれの最後の出力を連結したものが全結合層への入力となっている。全結合層は2つの値を出力し、その値を用いてSoftMax層にて、行動促進ツイートか否かを判別する。

### 3.3 BERTを用いた抽出手法

文脈を考慮できる方法の一つとしてBERT [17] が挙げられる。BERTはTransformerをベースとしている。TransformerはRNNやCNNを使わず、self-attentionのみを使用したモデルであり、長距離の依存関係を捉えることができる。そこで本論文では、文脈を考慮できるBERTを用いて行動促進ツイートの抽出を行う。

BERTは大規模な生コーパスでpre-trainingし、各タスクでfine-tuningするという2ステップからなっている。そのため、様々なタスクに対応でき、少ないデータでの転移学習も期待できる。本論文では、BERTの実装に機械学習ライブラリであるTensorflow<sup>3</sup>を使用する。使用するBERTのネットワークの構造を図3に示す。pre-trainingのモデルは、日本語のツイートに適用するにあたり、多言語モデルを利用することが考えられるが、解析単位がほぼ文字であり、単語単位のタスクに用いることには適さないと考えられる。そこで、解析単位が単語である日本語学習済みモデルの「BERT日本語Pretrainedモデル [18]」を用いる。特徴ベクトルは、ツイートに対して形態素解析を行い、BERT日本語pretrainedモデルに入力して、分散表現を行ったものを使用する。

BERTの各種ハイパーパラメータはグリッドサーチにより決定する。隠れ層はTransformer層が12、ベクトルサイズは768、アテンションのヘッド数は12、入力の最大系列長は90、バッチサイズは32、エポック数は10、学習率は0.00002、学習率のウォームアップ率は0.1、ドロップアウト率は0.1となっている。入力ユニット数はツイートに含まれる単語数で、SoftMax層にて行動促進ツイートか否かを判別する。

3: Tensorflow <https://www.tensorflow.org/>

表 2 提案手法による行動促進ツイート抽出結果

手法	適合率	再現率	F 値	AUC
LSTM	0.9418	0.8734	0.9064	0.9673
BLSTM	0.9257	0.9046	0.9150	0.9640
BERT	0.9305	0.9440	0.9372	0.9790

## 4 実 験

本論文では 2 種類の実験を行う。まず、提案した 3 つの手法による大規模災害時の行動促進ツイート抽出の比較実験を行う。次に、抽出した大規模災害時の行動促進ツイートと 2018 年の災害時の行動促進ツイートの比較実験を行い、分析する。

### 4.1 実験 1:提案手法の比較実験

提案した LSTM, BLSTM, BERT のいずれが、大規模災害時の行動促進ツイートの抽出に適しているかの比較実験を行う。  
実験データ

実験データは関東地方に甚大な被害をもたらした 2019 年の台風 15 号を対象とする。この台風の上陸前から上陸後の期間 (9 月 6 日～9 月 16 日) を対象に「台風」のクエリを用いて無作為に収集した 10,000 件のツイートをを用いる。収集したツイートに対し、アンケートにより行動促進ツイートであるか、そうでないかの判定を行う。アンケートにはクラウドソーシングを利用した。判定の方法は、1 ツイートあたり 5 人の被験者が判定する。3 人以上が行動促進であると答えたツイートを正例とし、全被験者が行動促進でないと答えたツイートを負例とする。その結果、正例 2,406 件、負例 4,795 件の合計 7,201 件のツイートが正解付きデータとして判定された。これらの正解付きデータのうち、正例、負例の数を同程度の数にするために、アンダーサンプリングを行う。具体的には、負例として判定された 4,795 件のツイートから無作為に抽出した 2,406 件のツイートを実験データの負例データとして利用する。その結果、使用する実験データは正例、負例ともに 2,406 件で合計 4,812 件である。

#### 実験方法

3 章で提案した 3 つの手法を用いて行動促進ツイートを抽出する。実験データの正例、負例からそれぞれ 20% のツイートを無作為に抽出し、テストデータとする。残りの 80% のツイートを教師データとし、提案した LSTM, BLSTM, BERT を用いてそれぞれ学習を行う。3 つの学習したモデルを用いてテストデータを判定し、適合率、再現率、F 値、Area Under the Curve(AUC) を求め、比較を行う。

#### 結果と考察

提案手法の適合率、再現率、F 値、AUC の結果を表 2 に示す。そして、3 つの手法それぞれの AUC のグラフを図 4 に示す。AUC の結果より、構築されたそれぞれの予測モデルの識別精度が良いため、大規模災害時の行動促進ツイート抽出に適しているといえる。また、これまで提案してきた LSTM の結果は、3 つの手法の中で適合率が最も高く、再現率と F 値が最も低い結果となっている。それに対して BERT では、再現率と F 値が最も高い結果を示している。

はじめに抽出結果の考察を行う。3 つの手法いずれも正しく抽出できた例を表 3 に示す。No.1 や No.2 のような、被災地周辺に関する被災者にとって有益な行動促進情報が正しく抽出できた。台風 15 号の大規模な被害が主に関東の局地的なところで生じていたために地域名がある程度限られていた点、及び被害が断水や停電、倒壊などに集中した点より、学習データとテストデータの内容が近いものが多く、3 つの手法とも正しく抽出できたと考えられる。しかしながら、No.3 のように、災害時にそのツイートを読んでも被災者にとって役に立たないような情報も抽出している。これは、用いたデータセットは「行動促進している」という判断の下、アンケートにより正例を決定したため、その正例の中には一般の被災者にとって役に立たない行動促進ツイートも多く含まれていた。そして、災害時における役に立たない行動促進ツイートは、No.3 のように内容が「お気をつけください」のみで、行動促進部分に関する説明が記載されていないものが多く存在していた。そのため、抽出精度が良かったと考えられる。したがって、行動促進ツイートが災害時において有益かどうか判定する必要があると考えられ、今後の課題である。

また、BERT のみが正しく判定できた例を表 4 に示す。No.1 と No.2 は BERT のみが抽出できた例で、投稿ユーザの経験から行動すべき情報を伝えている行動促進ツイートである。投稿ユーザの過去の経験から伝える行動促進情報は災害時において、具体的にどうすべきか記載されているため、被災者にとって有益だと考えられる。また、No.3 と No.4 は行動促進していないと判断された負例の例である。これらのツイートは、他人に行動を促進しておらず、ツイートをしている本人の意志表示をしているツイートである。ここでは BERT のみが行動促進ツイートでないと正しく判定できている。LSTM と BLSTM では誰に対しての行動なのかを考慮できていないため誤抽出してしまったと考えられる。

次に、3 つの手法の実行時間を比較する。本実験には GPU に Geforce GTX 1080Ti を用いた。3 つの手法の学習と評価それぞれ 1,000 回行った平均を求め、結果を表 5 に示す。結果から、学習にかかる時間は BERT が最も早い、実際のテストデータの計算時間は LSTM が早いという結果となった。しかしながら、合計を見ると BERT が最も早くなっていることがわかる。

以上の事より、3 つの手法とも災害時の行動促進ツイートにある程度抽出できているといえるが、精度と計算時間を鑑みると BERT が災害時の行動促進ツイート抽出に最も適しているといえる。

### 4.2 実験 2:大規模災害時とそれ以外の災害時の比較実験

大規模災害時の行動促進ツイートとそれ以外の災害時の行動促進ツイートとの比較実験を行い、それぞれの特徴分析を行う。大規模災害時の行動促進ツイートの対象の災害は、関東地方に甚大な被害をもたらした 2019 年の台風 15 号 (以下、台風 15 号) を用い、それ以外の災害には、2018 年に発生した台風 (以下、2018 年の台風) を対象とする。ここでは、2018 年に関西

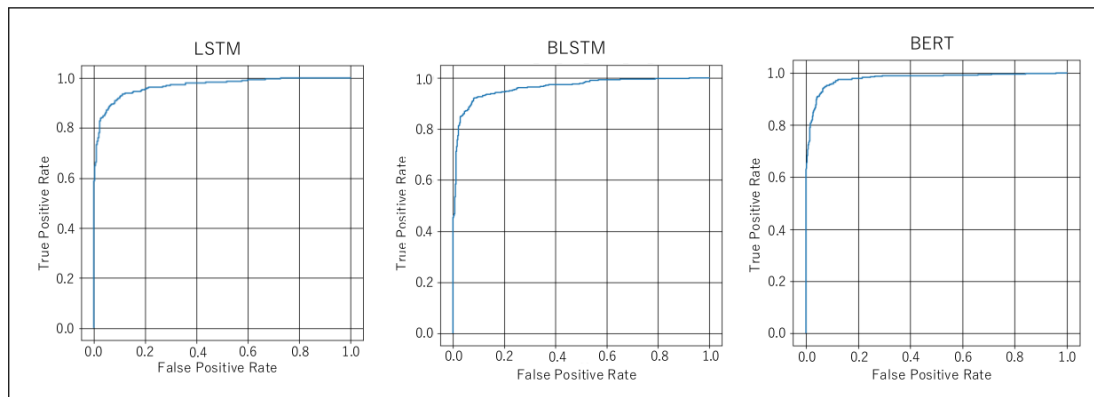


図 4 3つの手法の AUC

表 3 実験 1:3 つの手法いずれも抽出できた行動促進ツイートの例

No	正解	LSTM	BLSTM	BERT	ツイート
1	○	○	○	○	台風 15 号により被災された九十九里町民の方を対象に、停電が復旧するまでの間、以下の入浴施設が無料開放となります。タオル、シャンプーなどは各自でご準備ください。国民宿舎サンライズ九十九里日時:9 月 16 日 (月) 各日 15 時 20 時住…
2	○	○	○	○	台風 15 号は今後八丈町に最接近する見込みです雨、風が急に強まり波も高くなる事が予想されます不要不急の外出は控え海岸付近は近づかないでください。なお三根公民館、町役場研修室、榎立公民館、中之郷公民館、末吉公民館を自主避難所として開設しています避難…
3	○	○	○	○	関東住まいの皆さん台風は大丈夫でしょうか…今更のツイートですが、お気をつけください。今年は気候の変動が激しい…

表 4 実験 1:BERT のみ正しく判定できた行動促進ツイートの例

No	正解	LSTM	BLSTM	BERT	ツイート
1	○	×	×	○	去年の台風 21 号で屋根が飛ばされた経験で言うと、災害時に配られる薄いブルーシートは早ければ 3 ヶ月ほどで日光で劣化して穴が開くので、応急処置をした後に、厚めのブルーシートを上からも一枚覆って木材を打ってある程度線で押さえた方がいい。
2	○	×	×	○	台風ほんとと停電すると無力過ぎたのでゲーム機もモバイルバッテリーいっぱい充電したりしとくといよいよ…あと常温で食べられる食料と飲料水保管して冷凍庫の食料はなるべく減らしておいた方がいい…停電復旧に 2 日かかって冷蔵庫の食材が死んだ民より
3	×	○	○	×	さーてと、今日は早めに寝ようかなんか、連日寝不足だ台風が過ぎたタイミング、金曜は中秋の名月だし、揺れにちょっと気を付けようかなー。
4	×	○	○	×	9 日に車で遠出する用事があるが台風ピンポイントだな。。もう少し動向をみてみよう#台風 15 号 #天気

表 5 実験 1:提案手法のそれぞれのタスクにおける実行時間

手法	タスク	実行時間 (s)	合計 (s)
LSTM	学習	23.75445	29.59050
	評価	5.83605	
BLSTM	学習	34.27266	40.83368
	評価	6.56102	
BERT	学習	11.46488	19.57622
	評価	8.11134	

地方で甚大な被害をもたらした大規模な台風である台風 21 号のツイートは除いてある。災害時の行動促進ツイートの抽出には 3.3 節で提案した BERT を用いて抽出を行う。

#### 実験データ

大規模災害時の行動促進ツイートの実験データは実験 1 で用いたものを使用し、2018 年の災害時の行動促進ツイートの実験データはこれまでの研究 [1] で用いた台風のデータを使用する。

具体的には、2018 年の台風 (台風 21 号から 25 号を除く) のそれぞれの発生前から発生後に「台風」のクエリを用いて無作為に収集した 3,000 件のツイートを用いる。収集したツイートに対し、行動促進ツイートであるかどうかの判定を実験 1 と同様に、クラウドソーシングを用いてアンケートにより行う。判定の方法は 1 ツイートあたり 5 人の被験者が判定する。3 人以上が行動促進であると答えたツイートを正例とし、全被験者が行動促進でないと答えたツイートを負例とする。また、正例と負例の数を同程度の数にするために、アンダーサンプリングを行う。その結果、使用する 2018 年の災害時の行動促進ツイートの実験データは正例、負例ともに 556 件で合計 1,112 件である。

#### 結果と考察

表 6 にそれぞれの行動促進ツイートの例を示す。台風 15 号では関東周辺、特に千葉県が大きな被害を受けた。そして、被害が主に局地的なところで生じており、断水や停電、倒壊などに集中していた。そのため、No.1 のような、被災地周辺の状況や

表 6 実験 2:それぞれの台風時の行動促進ツイートの例

No	話題	ツイート
1	台風 15 号	台風一過、皆様ご無事ででしたでしょうか上野公園内では多くの木々が倒れました。コンクリートごと地面をえぐりながら倒れた大木もある程度、倒れた木々の対策は終わりましたが、まだ落ちている枝などございますのでお気をつけください…
2	台風 15 号	山武市消防防災課よりお知らせ (13:02 頃の放送) 飲料水の給水の再開をします。場所山武市役所松尾 IT 保健福祉センター時間 9/11(水)14:00~17:00 必ず容器を持参してください。#山武市#台風 15 号#断…
3	台風 15 号	今さらではありますが、去年の大阪台風の停電経験から申し上げると、最近の冷蔵庫はドアさえ開けなければ、3、4 日は中のもの大丈夫です。なので、食料等は常温保存してるものから消費して、冷蔵庫を開けるのは最後にした方がいいよ
4	2018 年の台風	台風の影響か曇り空ではありますが逆に湿気が多くて熱中症になりそうなのでお互い気を引き締めて暑さを乗り切りましょう
5	2018 年の台風	台風第 26 号イートゥー発生。今後の動向にご留意ください。

注意喚起をする行動促進ツイートや、No.2 のような、被災地の生活支援のための行動促進ツイートといった、テレビやラジオなどでは得難い被災地周辺の情報が多く見受けられた。また、No.3 のような、投稿ユーザの過去の経験から情報を伝える行動促進ツイートが見受けられた。これらの行動促進ツイートは大規模災害時において、被災者にとって有益な情報であると考えられる。また、行動促進部分に関しては、台風 15 号は 2018 年の台風と比べて、強い口調の行動促進が多く見受けられた。これは、台風 15 号がこれまでにないほど強い台風で、被害も大きかったため、行動を強く促しているツイートが多くなったと考えられる。そして、台風が上陸する前では予測による対策、台風上陸中では危険回避のための行動抑制、台風が去った後では被害状況に伴う注意喚起やアドバイスというように行動促進が異なっていることがわかった。

一方で 2018 年の台風では、No.4 のような弱い口調の行動促進が多く見受けられた。これら弱い口調の行動促進も抽出することができた。また、No.5 のような具体的にどのような行動を取るべきなのかが記載されていないツイートが多い傾向にあった。これらの行動促進ツイートが多かった理由として、2018 年の台風では台風発生時からあまり危機感をもっていなかったためだと考えられる。

これらの分析結果から、大規模災害時では、被災地周辺の状況やそれに伴う注意喚起をしている行動促進ツイートや被災地の生活支援のための行動促進ツイート、投稿ユーザの経験から情報を伝えている行動促進ツイートのような役に立つ情報を抽出し、提示する必要があると考えられる。

## 5 まとめと今後の課題

本論文では大規模災害時の行動促進ツイートの抽出手法を比較検討した。具体的には、大規模災害として 2019 年台風 15 号を対象とし、LSTM, BLSTM, BERT の 3 つの手法を提案した。大規模災害時の行動促進ツイート抽出の比較実験に行い、結果から 3 つの手法いずれも抽出精度が良く、提案した 3 つの手法が大規模災害時の行動促進ツイート抽出に適していることが確認できた。その中でも BERT が最も良い結果となった。また、抽出した大規模災害時の行動促進ツイートと 2018 年の災

害時の行動促進ツイートの比較分析から、それぞれの行動促進の表現や内容、傾向が異なることが確認できた。これら傾向の異なる行動促進ツイートも BERT により抽出可能であることもわかった。

今後の課題として暗示的行動促進ツイートの抽出やデマを含む行動促進ツイートの抽出、有益な行動促進ツイートの抽出や行動促進ツイートに対しての警告提示手法の提案が挙げられる。

## 謝 辞

本論文の一部は JSPS 科研費 17K00430, 18H03342, 19H04221, 19H04218, 及び、私学助成金（大学間連携研究補助金）の助成によるものである。ここに記して謹んで感謝の意を表する。

## 文 献

- [1] 米田 吉希, 見塚 圭一, 鈴木 優, 灘本 明代, “機械学習を用いた災害時行動促進情報抽出手法の比較検討”, WebDB Forum 2019 論文集, pp. 61–64, 2019.
- [2] A. Nadamoto, M. Miyabe and E. Aramaki, “Analysis of Microblog Rumors and Correction Texts for Disaster Situation”, Proc. of International Conference on Information Integration and Web-based Applications & Services, pp. 44–52, 2013.
- [3] 見塚 圭一, 米田 吉希, 鈴木 優, 灘本 明代, “災害時の行動促進ツイート抽出手法”, 第 11 回データ工学と情報マネジメントに関するフォーラム (DEIM2019), I7-4, 8 pages, 2019.
- [4] Z. Ashktorab, C. Brown, M. Nandi and A. Culotta, “Tweedr: Mining twitter to inform disaster response.”, Proceedings of the 11th International ISCRAM Conference, 2014.
- [5] 眞中 純美, 児玉 史緒里, 内田 理, 石井 啓之, 宇津 圭祐, “ニュース投稿をリツイートしたユーザの投稿に着目した災害関連情報の自動収集に向けた一検討”, 東海大学紀要. 情報通信学部 Proceedings of the School of Information and Telecommunication Engineering, Tokai University, Vol.9, No.2, pp.7-14, 2016.
- [6] K. Rudra, S. Ghosh, N. Ganguly, P. Goyal, and S. Ghosh, “Extracting situational information from microblogs during disaster events: a classification-summarization approach”, In *CIKM '15 Proceedings of the 24th ACM International on Conference on Information and Knowledge Management*, pp. 583–592, 2015.
- [7] 川崎 凌摩, 松下 光範, 宋 晨潔, 藤代 裕之, “Twitter からの救助要請の抽出と検証。—2018 年 7 月の西日本豪雨災害ツイートを対象として—”, 第 11 回データ工学と情報マネジメントに関する

- るフォーラム (DEIM2019), I7-5, 8 pages, 2019.
- [8] M. A. Cameron, R. Power, B. Robinson, and J. Yin, “Emergency situation awareness from twitter for crisis management”, In *WWW '12 Companion Proceedings of the 21st International Conference on World Wide Web*, pages 695–698, 2012.
  - [9] 湯沢 昭夫, 小林 亜樹, “マイクロブログにおける感動詞との共起を利用した検索語抽出システム”, 第 11 回データ工学と情報マネジメントに関するフォーラム (DEIM2019), C8-4, 8 pages, 2019.
  - [10] 北田 剛士, 風間 一洋, 榑 剛史, 鳥海 不二夫, 栗原 聡, 篠田 孝祐, 野田 五十樹, 齊藤 和巳, “東日本大震災時のツイートのトピック系列の可視化と分析”, 人工知能学会全国大会論文集, JSAI2015, 2B3NFC02a1-2B3NFC02a1, 2015.
  - [11] 佐藤 翔輔, 今村 文彦, “2018 年西日本豪雨災害における「#救助」ツイートの実態:2017 年 7 月九州北部豪雨災害との比較分析”, 自然災害科学, Vol.37 No.4 pages 383–396, 2019.
  - [12] 榎田 宗丈, 吉野 孝, “災害情報に着目したマイクロブログのつぶやき分析”, 2016 年度 情報処理学会関西支部 支部大会 講演論文集, G-14, 2016.
  - [13] S. Hochreiter and J. Schmidhuber, “Long Short-term Memory.”, In *Proceedings of the Neural Computation*, vol.9, pp. 1735–1780, 1997.
  - [14] M. Asahara, “NWJC2Vec: Word embedding dataset from ‘NINJAL Web Japanese Corpus’,” *Terminology: International Journal of Theoretical and Applied Issues in Specialized Communication*, vol. 24, no. 2, pp. 7–25, Feb. 2018.
  - [15] D. P. Kingma and J. Ba, “Adam: A Method for Stochastic Optimization”, *International Conference on Learning Representations*, 2014.
  - [16] Mike Schuster and Kuldip K Paliwal, “Bidirectional recurrent neural networks.”, *IEEE Transactions on Signal Processing*, Vol. 45, No. 11, pp. 2673–2681, 1997.
  - [17] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. “BERT: Pre-training of deep bidirectional transformers for language understanding.”, *arXiv preprint arXiv:1810.04805*, 2018.
  - [18] 柴田 知秀, 河原 大輔, 黒橋 禎夫, “BERT による日本語構文解析の精度向上”, 言語処理学会 第 25 回年次大会, pp.205-208, 2019.