

行動促進ツイートからの根拠部分抽出手法

米田 吉希[†] 見塚 圭一[†] 鈴木 優^{††} 灘本 明代^{†††}

[†] 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

^{††} 奈良先端科学技術大学院大学情報科学研究科 〒630-0192 奈良県生駒市高山町 8916-5

^{†††} 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: [†]{m1824003,m1724005}@s.konan-u.ac.jp, ^{††}ysuzuki@is.naist.jp, ^{†††}nadamoto@konan-u.ac.jp

あらまし SNS 上にはユーザに行動を促進する情報が多数存在する。その中でも Twitter は、人それぞれが様々な状況において投稿するため、Twitter 上にはユーザに対して多種多様な行動を促進するツイートがある。本研究ではそのようなツイートを行動促進ツイートと呼ぶ。行動促進ツイートの中でも、行動を促進する理由の説明のように、行動促進の根拠を含む行動促進ツイートが多数ある。しかしながら、根拠を含む行動促進ツイートは内容の真偽に関わらず、根拠を含まない行動促進ツイートよりも信用されやすいと考えられる。また、根拠の内容によってユーザの行動に与える影響が異なると考えられる。その為、行動促進ツイートの根拠を抽出する必要があると考えた。そこで本論文では、行動促進ツイートから根拠部分を抽出する手法を提案する。

キーワード Twitter, 情報抽出, 機械学習

1 はじめに

近年、Twitter や Facebook 等の SNS の普及により、誰もが簡単に情報の発信や収集を行うことが可能となっている。その中でも Twitter は、一般人や専門家等の様々なユーザが自由に内容を記述できるため、多種多様な内容の膨大な量のツイートが投稿されている。そして、ある話題に対するツイートの中には他人に行動を促進しているツイートが多数存在する。例えば「地震が発生したら、玄関などの扉を開けて出口の確保をしましょう。」のような行動を促進しているツイートがある。これまで我々はこのように他人に行動を促進しているツイートを行動促進ツイートと呼び、行動促進ツイートの抽出手法を提案してきた [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「揺れが収まってから火を消しましょう」のようにツイート内に明らかに行動を促進している部分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部分は存在しないが、他人に行動を促しているツイートである。例えば「地震で停電している方はブレーカーを落としておかないと、復旧時に火災の危険があります」というツイート内には、行動を促している部分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。

行動促進ツイートの中には、行動促進部分に対して、何故そのような行動をすべきなのか等の理由の説明が記述されているツイートが存在する。例えば、「地震でドア枠が歪んで開かなく

なるので、まず出入口は確保しましょう。」と「出入口は確保しましょう。」は同じ行動を促進をしている。前者は「地震でドア枠が歪んで開かなくなるので」という、「出入口は確保しましょう。」に対する理由の説明が記述されているが、後者は記述されていない。本研究ではこのような、行動を促進する理由の説明もしくは目的の説明、または行動促進部分によってもたらされる効果の説明を「根拠部分」と呼ぶ。

また、根拠部分を含む行動促進ツイートは内容の真偽に関わらず、根拠部分を含まない行動促進ツイートよりも信用されやすいと考えられる。その為、ユーザが行動促進ツイートの内容を深く考えずに拡散したり、実際に行動に移す可能性が高くなる。そこで我々は、行動促進ツイートに対してアラートを出すことにより、ユーザに内容を考える機会を提供できると考えた。また、根拠部分の内容によって行動促進ツイートの重要度が異なるため、考慮する必要があると考えられる。例えば、「夜中にあった地震は本震ではないと気象庁が言っています。まだ余震もありますので、気をつけてください」と「まだ余震もありますので、気をつけてください」では同じ行動促進の内容であると共に、それぞれ根拠部分を含んでいる。しかしながら、前者は後者よりも根拠部分が具体的な内容であるため、前者のほうがユーザが判断するのにより役に立つと考えられる。また、根拠部分には、投稿者の経験に基づいた内容や「～かもしれない」というような不確かな内容等があり、それぞれがユーザの判断に与える影響が異なる。したがって、根拠部分の有無だけでなく、その内容も考慮することでユーザの判断の支援ができると考えられる。そこで本研究では、根拠部分の特徴語に着目した、根拠部分を含む行動促進ツイートの抽出手法を提案する。そして、その行動促進ツイートに対して、根拠部分の有無や内容による重要度を考慮したアラートの提示を目的とする。本研究により、誤った情報を鵜呑みにし、不利益を被るユーザを減少させられると考えられる。本研究の全体の流れを以下に示す。

表 1 根拠部分の例

地震はいつ来るかわからないので備えをしておきましょう。
地震、余震がきた時のためにできる限りの備えをしておきましょう。
最優先に守りたいのは頭部です。落下物から守るために、カバンでも厚めの雑誌でもいいから頭にのせて、ガードしてください。そして机の下に避難すること！緊急速報後の 10 秒ではこの程度が限界ですが、頭を守ることで生命の危機は大半回避できます。
地震の際不足するのは飲料水ではなく、トイレを流す水です。水が出るうちに浴槽等に溜めておくのを勧めます。
地震の被害ひどいですね。避難される方も多いでしょうが、避難の際は戸締りと一緒に水道とガスの元栓のチェック、それと“電気ブレーカーを落とす”のも忘れないで下さいね。電気止まっているので、通電した際の火災防止だそうです。ブレーカーは忘れがちなので気をつけねばです。

- (1) Twitter からある話題に関するツイートの取得
- (2) 行動促進ツイートの抽出
- (3) 根拠部分を含む行動促進ツイートの抽出
- (4) 根拠部分の抽出
- (5) 根拠部分を含む行動促進ツイートの重要度の算出
- (6) 行動促進ツイートに対してアラートの提示

本論文では、(3) の行動促進ツイートから根拠部分を含む行動促進ツイートを抽出する手法を提案する。具体的には、根拠部分を含む行動促進ツイートと含まない行動促進ツイートとの比較および分析の結果、根拠部分に特徴が見られた。その特徴を用い、(2) で抽出した行動促進ツイートから、ルールベースにより根拠部分を含む行動促進ツイートを抽出し、評価実験と考察を行う。

以下、本論文では 2 章で関連研究について述べ、3 章で根拠部分の定義について述べる。そして、4 章で根拠部分を含む行動促進ツイートを抽出する手法について述べ、5 章で提案手法の評価実験を行い、その考察について述べる。最後に 6 章で本論文のまとめと今後の課題について述べる。

2 関連研究

文章からの根拠部分の抽出の研究は数多くされている。飯田ら [2] は文章中の意見や命題などの根拠を抽出するために、機械学習を用いて根拠-帰結関係を抽出する手法を提案しており、5 分割交差検定での評価により、提案手法が有用であることを示している。藤川ら [3] は情報の真偽の判断を支援するために、Twitter 上の流言を疑いの有無と根拠の有無の二つの観点から信疑の分類を SVM を用いて行い、提案手法の有効性を示している。福永ら [4] はタスク指向対話システムにおいて、SVM を用いた根拠となる単語の抽出手法を提案しており、不動産データにおいて提案手法の有意性を示している。Lei ら [5] は商品レビューから各評価項目の評価の根拠を抽出する手法を提案している。これらの研究は根拠という点では類似しているが、本研究では、Twitter 上からある話題に関する行動促進ツイートを対象とし、その行動促進ツイートから根拠部分を抽出するという点で異なる。

また、本研究では根拠部分の抽出のため因果関係にも着目しているが、文章中の因果関係に着目した研究も多く行われている。坂地ら [6] は文中の 3 つの表現の位置関係の分析を行い、新聞記事から因果関係を自動的に抽出する手法を提案している。

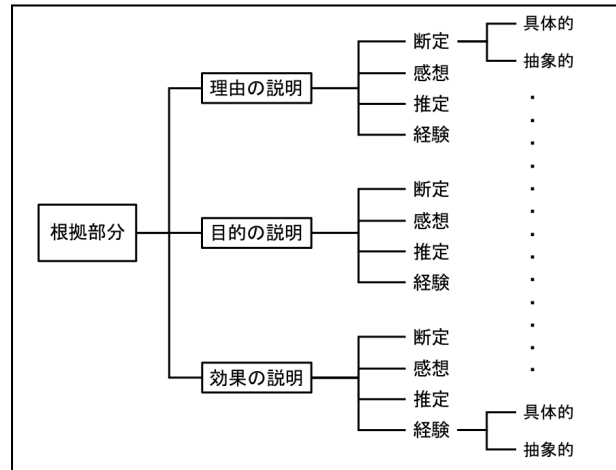


図 1 根拠部分のタイプ

乾ら [7] は接続標識「ため」を含む複文から、獲得可能な因果関係を自動的に分類する手法を提案し、分析と実験より有用性を示している。高野ら [8] は評判情報に含まれる因果関係を用いて、評価理由の抽出手法を提案している。これらの研究は、文章中から手がかりとなる語を用いて原因や理由を抽出するという点では類似しているが、本研究では、原因や理由の部分の機能や出現する名詞の頻度に着目して抽出するという点で異なる。

3 行動促進ツイートの根拠部分の定義

行動促進ツイートの根拠部分はいくつかのタイプに分類される。図 1 に我々が分類した根拠部分のタイプを示す。根拠部分は、行動を促進する理由の説明もしくは目的の説明、または行動促進部分によってもたらされる効果の説明の 3 種類に分類され、根拠部分を含む行動促進ツイートの例を表 1 に示す。行動を促進する理由の説明とは、例えば「エレベーターは閉じ込められる危険があるため、エレベーターを使わないようにしましょう。」という行動促進ツイートの場合、「エレベーターは閉じ込められる危険があるため、」が「エレベーターを使わないようにしましょう」という行動促進に対する理由の説明である。行動を促進する目的の説明とは、例えば「エレベーターで閉じ込められた時は救助を求めるために、まずは備え付けのインターホンを利用しましょう。」という行動促進ツイートの場合、「救助を求めるために、」が「インターホンを利用しましょう。」という行動促進に対する目的の説明である。また、行動促進部分によつ

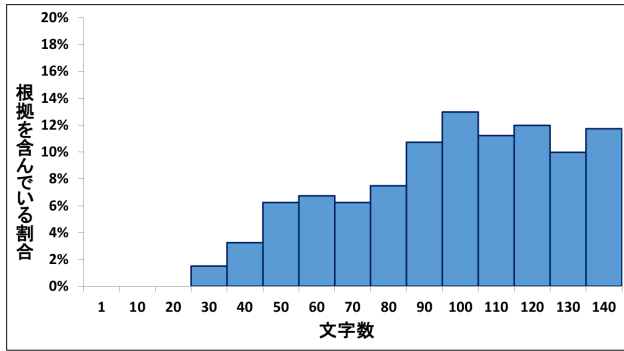


図 2 根拠部分を含む行動促進ツイート数と文字数との関係

表 2 手がかり語

ので	から	ため	なので	だから
すれば	すると	することで	することにより	

てもたらされる効果の説明とは、例えば「エレベーター内で地震に襲われたら、階のボタンを全部押してください。そうすることで、どこかの階で停止してくれる可能性が高くなります。」という行動促進ツイートの場合、「そうすることで、どこかの階で停止してくれる可能性が高くなります。」が「階のボタンを全部押してください。」という行動促進に対する効果の説明である。

また、図 1 に示すように根拠部分の中には、推定や感想といった機能的に働く根拠が存在する。その機能によって、それぞれの行動促進ツイートの重要度が異なるため、ユーザの判断の支援になり、重要であると考えられる。

根拠部分は図 1 に示すタイプに分類されるが、本論文ではタイプを考慮せず、行動促進ツイートから根拠部分を含む行動促進ツイートの抽出を目的とする。

4 行動促進ツイートからの根拠部分の抽出

4.1 根拠部分を含む行動促進ツイートと文字数の関係

行動促進ツイートの中でも、根拠部分を含む行動促進ツイートは根拠部分を含まない行動促進ツイートよりも文字数が多くなる場合がある。そこで、行動促進ツイートの文字数が根拠部分の有無に影響すると考え、実際に根拠部分を含む行動促進ツイートと文字数との関係を調べた。行動促進ツイート 556 ツイートを分析した結果、図 2 のような結果になった。これにより、文字数が多いほど根拠部分を含むツイートの割合は高くなっているが、文字数によって根拠部分の抽出が一概に出来るとは言いがたいことがわかった。

そこで本論文では、根拠部分を含む行動促進ツイートを自動的に抽出するために、ルールベースによって抽出する手法を提案する。

4.2 行動促進ツイートの抽出

これまで我々は、行動促進ツイートの抽出手法について研究してきた [1]. [1] より、LSTM [9] の手法が行動促進ツイートを抽出する手法として最も適していることがわかったため、LSTM

を本研究でも用いる。

LSTM を用いた行動促進ツイート抽出手法

行動促進ツイートの抽出時に、ツイート内の単語の出現順序を考慮し、学習させることにより、正しく判定が行えると考え、文の単語の順番を考慮することの出来る LSTM を用いて学習を行う。LSTM の実装には、Python の機械学習ライブラリである Chainer¹ を用いる。

素性は、ツイート内に出現する単語を Word2Vec を用いてベクトル化したものを使用する。使用する単語は、教師データ内から url と Twitter のユーザ名を取り除いた後のデータ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、文の単語の流れを捉えることにより、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。LSTM の各種パラメータは、隠れ層の数は 1、ユニット数は 200、バッチサイズは 250、エポック数は 300、学習率は 0.001、オプティマイザーは Adam を利用する。

4.3 ルールベースを用いた根拠部分の抽出手法

行動促進ツイートを抽出した後、その中から根拠部分を含む行動促進ツイートを抽出する。根拠部分には「多分電気も水道も止まるので、冷蔵庫や冷凍庫のものは、悪くなりそうなものから早めに食べてください！」や「大地震はいつどこで起きるか予測不可能です。だからいつ来てもいいように備えましょう！」といったように「ので」や「だから」のような、根拠部分の手がかりになる特徴語がツイート内に現れやすい。本論文ではこの特徴語を「手がかり語」と呼ぶ。そこで我々は、556 件の行動促進ツイートを分析し、表 2 に示すように手がかり語 9 語を決定した。

しかしながら、根拠部分を含む行動促進ツイートに手がかり語が存在しないツイートもある。例えば、「避難する前にブレーカーを落として避難しましょう！通電火災になると大変です。」は根拠部分を直接的に表現していないが、「通電火災になると大変です。」が「避難する前にブレーカーを落として避難しましょう！」という行動を促進する理由であると解釈される場合がある。そこで 1,176 の行動促進ツイートの分析をし、根拠部分を含むツイートと含まないツイートの名詞の出現単語に差があることがわかった。根拠部分を含むツイートにはある話題に関連する名詞が存在し、根拠部分を含まないツイートにはその名詞が存在しないことがある。例えば、地震の話題に関する行動促進ツイートで、根拠部分を含むツイートには「火災」や「津波」といった、地震に関連する名詞が多く存在しているが、根拠部分を含まないツイートにはあまり存在していない。しかしながら、ある話題に関連する名詞で、関連度が高い名詞だと根拠部分を含むツイートにも根拠部分を含まないツイートにも存在する。例えば、「災害」や「余震」はどちらのツイートにも多く含まれるため、これらは特徴語としてあまり適していない。また、ある名詞の出現回数が少ないと、その名詞はその話題に

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

表 3 ダイエットの関連語

代謝	燃焼	ホルモン	エネルギー	ビタミン	基礎	吸収
不足	たんぱく	分泌	促進	脳	刺激	内臓
バナナ	値	作用	消化	血糖	分解	成分
体温	糖	血液	中枢	成長		

表 4 地震の関連語

水	避難	場合	火災	ブレーカー	携帯	通電
物	準備	バッテリー	津波	規模	食料	懐中
電灯	トイレ	保険	備蓄	うち	家具	落下
ガラス	機	火	電話	電池	風呂	

あまり関連していないと考えられるため、特徴語として適しているとは言えない。そこで、根拠部分を含むツイートと含まないツイートに含まれる名詞の TF 値を比較した。その結果、根拠部分を含むツイートに出現する名詞の TF 値を根拠部分を含むツイートに出現する名詞の TF 値で割った値が 0.25 以下かつ、根拠部分を含むツイートでの出現回数が 16 回以上の名詞をその話題の関連語として決定する。実際に、ダイエットの話題と地震の話題についての関連語を抽出した。ダイエットの話題は 1,080 ツイートを分析し、26 語の関連語を抽出した。また、地震の話題は 1,176 ツイートを分析し、27 語の関連語を抽出した。その結果を表 3 と表 4 に示す。

そして、根拠部分を含む行動促進ツイートを抽出するルールは、行動促進ツイート内に手がかり語もしくは話題の関連語を 1 語以上含むツイートを、根拠部分を含む行動促進ツイートとして抽出する。

また、「本震があるかもしれないので皆さん色々と備えましょう!」の「かもしれない」といった、根拠部分が推定として働く表現や「まだ余震があると思うので気をつけましょう。」の「思う」といった、根拠部分が感想として働く表現等がある。そこで我々は、「かもしれない」や「思う」といった機能表現に着目することにより、根拠部分の抽出を行うことができると考えた。この機能表現とは全体として 1 つの機能的な意味を持つ表現のことである。本研究では、宇津呂らの提案する機能表現 [10] を用いて抽出する。例えば、「僕は野球がしたい。」や「晴れているなら野球をする。」という文のように「たい」や「なら」といった、文に対して願望や順接などの意味を付与する表現である。この機能表現を用いる際には、松吉ら [11] の提案する機能表現辞書つつじを使用する。本研究において、根拠部分の機能は断定、感想、推定、経験の 4 つを用い、機能における条件は表 5 に示す。

5 実験

行動促進ツイートから根拠部分を含む行動促進ツイートを抽出する手法の有用性を調べるために、ルールベースを用いた抽出の実験を行った。

5.1 実験データ

実験データはアンケートにより、行動促進ツイートに根拠部分が含まれているかどうかを手手で判定したツイートをを用い

表 5 根拠部分の機能における条件一覧

条件
断定
手がかり語の直前が「だ」や「である」のような「断定」の機能表現
感想
手がかり語の直前が「思う」や「感じる」のような「感想」の機能表現
推定
手がかり語の直前が「みたい」や「らしい」のような「推定」の機能表現
経験
手がかり語の直前が過去形

表 6 抽出結果

話題	適合率	再現率	F 値
ダイエット	0.750	0.791	0.769
地震	0.803	0.815	0.808

る。アンケートにはクラウドソーシングを利用した。具体的には、“ダイエット”と“地震”の 2 つの話題に関する行動促進ツイートをを用いた。ダイエットに関する行動促進ツイート 1,172 ツイート、地震の行動促進ツイート 2,174 ツイートに対し、5 人の被験者によって根拠部分を含むか含まないかを判定した。5 人中 3 人以上が根拠部分を含むと答えたツイートを正例とし、同様に 3 人以上が根拠部分を含まないと答えたツイートを負例とする。その結果、ダイエットの話題に関して正例 540 ツイート、負例 540 ツイートの合計 1,080 ツイート、地震の話題に関して正例 588 ツイート、負例 588 ツイートの合計 1,176 ツイートの行動促進ツイートを実験データとして決定した。ここでダイエットの話題を選択した理由はツイートをしている人が健康についての行動を促進する傾向があると考えたためである。また、地震の話題を選択した理由は、災害時は活発に行動促進ツイートが行われ、正確な判断が出来ない状況のため、ツイートが信じられやすいと考えたためである。

5.2 実験方法

ルールベースによる根拠部分を含む行動促進ツイートの抽出精度を調べるために交差検定を用い、分割数は 5 とした。話題の関連語は表 3 と表 4 に示すように、ダイエット 26 語、地震 27 語を用いた。そして、各話題の実験データ内に手がかり語もしくは話題の関連語を 1 語以上含むツイートを、根拠部分を含む行動促進ツイートとして抽出した。5 分割交差検定による適合率、再現率、F 値を話題ごとにそれぞれ求めた。

5.3 結果と考察

表 6 に 2 つの話題の適合率、再現率、F 値を示す。また表 7、表 8 にそれぞれの話題の抽出結果の一例を示す。実験結果から、ダイエットと地震のどちらの話題に対しても、ルールベースによって根拠部分を含む行動促進ツイートの抽出がある程度出来ていることがわかる。これは表 7 と表 8 の No.1 や No.2, No.5 の例のように、行動促進ツイートの根拠部分には直接的に表現している手がかり語や、その話題に関連のある語が含まれているツイートが多く存在する為だと考えられる。しかしながら、

表 7 ダイエットの話題における抽出結果例

No.	正解	ルール	ツイート
1	○	○	ガムは1粒 10kcal にも満たない低カロリーで、適度な糖分と、たくさん噛むことで満腹中枢を刺激することができます。お腹が空いたらガムを噛む習慣をつけましょう。 #ダイエット
2	○	○	ソフトクリームみたいに、かぶりつくタイプのアイスでも、スプーンでちょっとずつ取りながら食べた方が、満腹中枢が刺激されて、より満腹感が得られます。スプーンで食べられるアイスなら、スプーンをもらうようにしましょう。 #ダイエット Mar 262016
3	○	×	乾物を食べよう！乾物は、水分が抜けることによって、栄養素が凝縮され、栄養価がアップ！乾物でダイエット食品と言えば寒天が有名ですが、切干大根、干しわかめ、高野豆腐、ひじきなどは、全て低カロリーで、植物繊維、カルシウム、鉄分などのミネラルが豊富。是非、食事に取り入れて。
4	×	○	何のために痩せるのか、目標や理由をまず見つけましょう。それだけで気持ちが変わります。 #ダイエット #美容
5	×	×	ダイエット目的の方は加圧トレーニングをしたあと、最低2時間は食事は摂らないでください。筋肉をつけたい方はヘルシーなお肉・お魚・豆製品とサラダを食べてください。

表 8 地震の話題における抽出結果例

No.	正解	ルール	ツイート
1	○	○	地震というものとは予測が不可能だ。従って地震が～時に来るとか言うのは嘘の可能性が非常に高いので無視するといひ。
2	○	○	【札幌市地震関連情報】停電の影響により、自動火災報知設備のバッテリー切れの警報音を火災と間違えて通報し、消防隊が出勤する事案が多く発生しています。通報する際は、今一度よくご確認いただくよう、お願いいたします。
3	○	×	地震とわかったらすぐに布団を被りましょう。ガラスの破片や家具の角が直接当たらないで済みます冬だとあたたかいです
4	×	○	あとコレから夜に備えて、もし寝れる状況であれば枕元に靴・必要最低限の貴重品を置き、いざと言う時外に出れる格好で寝てください。暗いなか次の地震が怖くてスマホ見てしまいがちになりマスが、情報収集はラジオ等で、もし無ければ1~2時間おきに30分程度…と決めて利用した方が良いデス。
5	×	×	【北海道での震度6強の地震に関して】今朝の地震で被災された皆さま、まずは、安全確保にご留意され、二次災害の防止に努めてください。余震にはくれぐれもお気をつけください。被災地の皆さまのご無事・安全と、少しでも早い復興を心よりお祈り申し上げます。

抽出できなかった根拠を含む行動促進ツイートの例として、表7のNo.3と表8のNo.3が挙げられる。これらのツイートは手がかり語も話題の関連語も含まないために抽出できなかった。表7のNo.3のツイートでは、「乾物」や「寒天」がダイエットという話題に関連のある名詞であるにもかかわらず、実験データ内での出現回数が少ないために話題の関連語に含まれていない。その為、行動促進ツイートで根拠部分を含むツイートと根拠部分を含まないツイートとの出現単語の比較のみによる話題の関連語の決定では取りこぼしが多くあるため、他の手法による話題の関連語の追加や他のルールを設ける必要があると考えられる。さらに、適切な話題の関連語を決定するためにデータ数を増やす必要もある。また、表8のNo.3のツイートでは、根拠部分を直接的に表現していないが、このツイートを読んで「ガラスの破片や家具の角が直接当たらないで済みます。」が「布団を被りましょう。」という行動を促進する理由であると解釈される為であると考えられる。その為、手がかり語や話題の関連語のような単語の有無のみでの判断だけでなく、文脈関係を考慮する必要があることがわかった。

また、誤抽出の例として表7のNo.4と表8のNo.4が挙げられる。これらは手がかり語もしくは話題の関連語の有無のみで抽出しているため誤抽出された。その為、手がかり語や話題の関連語の再考やルールの追加、単語だけでなく文と文との関係も考慮する必要があると考えられる。

6 まとめと今後の課題

本論文では、ルールベースを用いた行動促進ツイートからの根拠部分を含む行動促進ツイートの抽出手法を提案した。具体的には、根拠部分の手がかり語9語を決定した。さらにある話題の関連語を、根拠部分を含む行動促進ツイートと含まない行動促進ツイートの出現単語のTF値を比較して決定した。そして、手がかり語と話題の関連語を用いて、行動促進ツイートから根拠部分を含む行動促進ツイートを抽出する手法の提案を行った。また、実験を行い、提案手法の有用性を測った。抽出の精度はある程度良い結果を得られた。しかしながら、抽出精度の向上の為に手がかり語と話題の関連語の決定方法の見直しやルールの追加が必要であることがわかった。さらに、話題の関連語で適切な名詞を決定するためにはデータ数を増やす必要がある。他にも、手がかり語と話題の関連語のような単語の有無のみでの判断だけでなく、文脈関係を考慮する必要があることがわかった。

今後の課題として、根拠部分抽出の精度を上げるために、機械学習等による別の抽出手法を考える必要がある。また、本論文では2つの話題のみを用いて実験を行っているが、他の話題を用いて実験を行い、他の話題にも適用できるか確認する必要がある。さらに、機能表現の利用や具体性に着目して、根拠部分を含む行動促進ツイートの重要度を算出する手法を考える必

要がある。

謝 辞

本論文の一部は JSPS 科研費 17K00430, 16K07973, 18H03342 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] 見塚 圭一, 鈴木 優, 灘本 明代, “Twitter 上の行動促進ツイート抽出手法の比較検討”, 第 11 回 Web とデータベースに関するフォーラム (WebDB Forum 2018), 信学技報, vol. 118, no. 213, DE2018-17, pp. 49-54, 2018.
- [2] 飯田 龍, 乾健太郎, 松本祐治, “根拠情報抽出の課題設計と予備実験”, 言語処理学会第 15 回年次大会, pp.817-820. 2009.
- [3] 藤川 智英, 鍛冶 伸裕, 吉永 直樹, 喜連川 優, “マイクロブログ上の流言に対するユーザの態度の分類”, 電子情報通信学会技術研究報告, DE, データ工学, pp.55-60, 2011.
- [4] 福永 隼也, 西川 仁, 徳永 健伸, 横野 光, 高橋 哲朗, “タスク指向対話における発話の分類と根拠抽出”, 言語処理学会第 24 回年次大会, pp.1035-1038, 2018.
- [5] Tao Lei, Regina Barzilay, and Tommi Jaakkola. Rationalizing Neural Predictions. In Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing, pp.107-117, 2016.
- [6] 坂地 泰紀, 増山 繁, “テキストマイニングによる因果関係抽出”, 第 54 回自動制御連合講演会, pp.456-459, 2011.
- [7] 乾 孝司, 乾 健太郎, 松本 裕治, “接続標識「ため」に基づく文書集合からの因果関係知識の自動獲得”, 情報処理学会論文誌 45(3), pp.919-933, 2004.
- [8] 高野 敦子, 池奥 渉太, 北村 泰彦, “因果関係に着目した口コミ Web サイトからの評価表現抽出”, 人工知能学会論文誌 24 巻 3 号 C, pp.322-332, 2009.
- [9] Sepp Hochreiter and Jurgen Schmidhuber, Long Short-term Memory. In Proceedings of the Neural Computation, Vol. 9, No. 8, pp.1735-1780, 1997.
- [10] 宇津呂 武仁, 松吉 俊, 土屋 雅稔, 鈴木 敬文, 島内 蘭, “自然言語処理における日本語機能表現の解析”, 語彙・辞書研究会 第 38 回発表会 資料集, pp. 1-8, November 2010.
- [11] 松吉 俊, 佐藤 理史, 宇津呂 武仁, “日本語機能表現辞書の編纂”, 自然言語処理, Vol. 14, No. 5, pp. 123-146, 2007.