

BERT と SNS を用いたコロナ渦における医師・専門家・知事から発せられたツイートに対する反応の分析*

柿崎 有紀[†]・酒井 由紀子^{††}・島内 宏和^{†††}

Reponses to tweets from physicians, experts, and governors during Covid-19 pandemic: An analysis using BERT

YUKI KAKIZAKI, YUKIKO SAKAI AND HIROKAZU SHIMAUCHI

ABSTRACT

In this study, the responders to the tweets during the Covid-19 pandemic from physicians, experts, and governors were analyzed using semantic orientations of words and neural language models in terms of persuasive health communication. A binary classification model was constructed that identifies whether a given tweet indicates the user was going out, as a fine-tuned BERT (Bidirectional Encoder Representations from Transformers). The constructed model was used to examine the increase of number of going out before and after the declaration of the state of emergency. From the results of the behavioral analysis, it is speculated that persuasive communication may have been more effective from physicians, experts, and governors in this order and the behavioral changes have been observed. In addition, a correlation between the results of emotion analysis using the word-emotion polarity correspondence table and behavioral changes was also observed. A limitation of this study is that the content of the tweets is not specifically analyzed.

Key Words: BERT, SNS, Twitter, COVID-19, Health Communication, Health Behavior

キーワード: BERT, SNS, Twitter, COVID-19, ヘルスコミュニケーション, 健康行動

1. まえがき

世界を揺るがしている新型コロナウイルス感染症 (Covid-19) は、日本においても 2020 年に始まる感染拡大が 2022 年 2 月現在で第 6 波を数え、いまだに収束の兆しが見えない。そのため、パンデミックにおけるヘルス・リスク・コミュニケーションのあり方が問われており、関連研究が盛んに行われている。たとえば、Berg らは保健機関による情報提供にかかわる論文を、2009 年以降の

* 令和 4 年 9 月 25 日 受付

令和 5 年 1 月 23 日 受理

[†] 工学部システム情報工学科・4 年 (原稿執筆時)

^{††} 帝京大学 共通教育センター・准教授

^{†††} 工学部工学科・大学院工学研究科・准教授

新型インフルエンザ (Swine flu) および Covid-19 に関して 1,053 件収集し、基準に合致した実証研究 48 件についてレビューを行っている。そのうち 33 件が Covid-19 に関するものである [1]。

パンデミックにおいては特に、「誰が」「何を」「どう伝える」ことで一般の人々の行動変容が期待できるか、すなわち説得的コミュニケーションは、感染拡大対策における大きな関心事である。本研究では、奥原剛ら[2]の実験研究の結果を受け、「医師」「専門家」「知事」のツイートとそれらのツイートに対するリプライと、単語感情極性対応表およびニューラルネットワークによる言語モデルを用いて、メッセージの受け手の反応について「感情」「行動」の2側面から分析し、行動変容に関する知見を得ることを目的とする。

2. 関連研究

本研究で追従する先行研究[2]を含む Covid-19 における説得的コミュニケーション研究、および Covid-19 におけるソーシャルメディアを用いた研究について選択的にふれる。

2.1 説得的コミュニケーション研究

行動変容に結びつくと考えられる説得的コミュニケーションの要素として、「誰が」「何を」「どう伝える」かがある。中でも「誰が」はその筆頭にあげられるが、それは送り手に対する「好意」と送り手の「権威」および「信憑性」が説得に貢献するからとされている。石川によれば、「信憑性」はさらに「専門性」と「信頼性」にブレークダウンすることができる[3, p.61]。Berg らのレビューによると、パンデミックの状況では「医師」の「信頼性」は高いが、政府の情報源については結果が安定していない[1]。Ricou らによる Covid-19 の情報提供を含む活動に関して、「誰」を信頼できるかを職種別に尋ねたポルトガルの研究では、医師、看護師、次いで科学者や教育者への信頼度が高く、政治家とジャーナリストへの信頼度が低いことが示されている[4]。

説得的コミュニケーションにおいてはもちろん、「何を」、すなわちメッセージそのものも重要な要素である。奥原らは「誰が」「何を」伝えれば、Covid-19 感染拡大抑制のために外出自粛を促すことができるのかについて、日本の成人 1,980 人の参加者を対象に 2020 年 5 月にランダム化比較実験研究を行った[2]。実験では、知事、公衆衛生の専門家、医師、患者、感染爆発地域の住民を情報発信者とした典型的な外出自粛を呼びかけるメッセージを作成した。参加者にはそのいずれかを読んでもらい、前後の「外出自粛の気持ち」の度合いを測定している。その結果、「医師」によるメッセージを読んだ人の方がほかのメッセージを読んだ人よりも外出自粛の気持ちが高くなった。その理由については、奥原らは「医師」の医療のひっ迫感をうったえる内容のメッセージが、「専門家」や「知事」の理性的なメッセージよりも、一般の多くの人々が持っている動物的に反応する心に強く刺さったのでであると、認知機能の二重過程理論をもとに解釈している[5]。

2.2 ソーシャルメディア分析

Berg らのレビューによると、Covid19 の情報提供に関する実証研究 33 件のうち 8 件がソーシャルメディアを対象としている。すべてがメッセージの内容と、再伝送や「いいね！」等の数による

エンゲージメントについて分析しているものである。外出自粛等が含まれる自己防衛行動を扱った研究が見当たらないため、今後取り組む必要があると主張されている[1]。

日本においても Covid-19 に関するソーシャルメディアの研究は行われている。Twitter のメッセージを対象としたものでは、感染状況と感情の変化を分析した研究がある[6,7]。また説得的コミュニケーションの「誰が」に係る研究としては、ワクチン接種に対するスタンスの変化とその要因を探った研究がある[8]。この研究では、手動によるアノテーションを用いて分類器を構築し、ツイートをワクチン接種に対するスタンスによって分類して、その変化と参照先を分析している。その結果、中立派から親ワクチン派に変化したユーザーに参照されたユーザーに医師、メディア、国の公式アカウントが多かったことを報告している。しかしながら、実際の行動変容への影響に直結した研究はやはり見当たらない。

3. 本研究の目的

「誰が」「何を」「どう伝える」ことで一般の人々の行動変容が期待できるかについて、直接扱った研究は少ない。その中で、奥原らの実験研究 [2]では、「誰が」だけでなく「何を」を扱っており、また外出自粛に対する影響までをエビデンスレベルの高いランダム化比較試験で測定している点が評価されている。しかしながら、測定しているのは「気持ち」の度合いであり、実際の行動ではない。また、一般的に、パンデミックの状況におけるメッセージの影響としての行動変容を実測することは容易ではない。

そこで、本研究では、Twitter 上の「医師」「専門家」「知事」のツイートとそれに対するリプライを行ったユーザーのツイートを分析することにより、三者から発せられたメッセージに対する受け手の「感情」への影響と、受け手が実際に外出を自粛したかどうか「行動」の変化に関する知見を得ることを目的とする。

4. 研究方法

「医師」「専門家」「知事」から 2021 年 7 月 5 日から 9 月 30 日の間に発せられたツイートとそれに対するリプライを分析の対象とした。2021 年 7 月 5 日は東京都に 4 回目の緊急事態宣言が発令された 7 月 12 日の一週間前にあたり、9 月 30 日は緊急事態宣言が解除された日である。三者から発せられたメッセージに対する受け手の「感情」については、東京都の 4 回目の緊急事態宣言の発令から解除までの期間における各リプライの形態素解析および単語感情極性対応表を用いて、「医師」「専門家」「知事」に対するものの感情値の平均を算出し比較する。

「行動」の分析については、深層学習を用いた自然言語処理モデルである BERT (Bidirectional Encoder Representations from Transformers, [9])を用いて各ツイートが外出したことを表すツイートかどうかを判別する識別器を構築する。識別器によりリプライを行ったユーザーの緊急事態宣言前後等の外出の頻度を推測し、実際の外出行動の変化について分析する。

4. 単語感情極性対応表による感情の分析

本章では、単語感情極性対応表を用いた「医師」「専門家」「知事」のツイートに対するリプライの感情分析について詳述する。

4.1 対象アカウントの選定

奥原らの実験研究[2]では、ランダム化比較実験において「医師」「専門家」「知事」と「患者」「感染爆発地域の住民」から発せられたメッセージを検証の対象としていた。本研究では Twitter 上のアカウントを用いるが、「患者」「感染爆発地域の住民」の属性についてはユーザープロフィール等からの推定は容易ではないことから、対象からは除外した。

「医師」については Twitter 上で「内科医」をキーワードとしてアカウントの検索を行い、そのうちフォロワー数が 100 人以上であったアカウントを対象とした。「専門家」は、アカウントのプロフィール欄において自身の所属を明らかにしており、かつ感染症の専門家と記載のあるものとした。「知事」のアカウントは、存在する現職知事のアカウントを対象とした。収集されたアカウントは「医師」422 件、「専門家」20 件、「知事」28 件、合計 470 件である。

4.2 感情分析に用いるツイートの収集

Twitter Standard API の Get Tweet timelines [10]を用いて「医師」「専門家」「知事」の 2021 年 7 月 5 日から 9 月 30 日までのツイートの収集を行った。収集された「医師」「専門家」「知事」のツイート数は医師が 62,280 件、専門家が 16,723 件、知事が 2,405 件である。

Twitter Standard API を用いて、「医師」「専門家」「知事」の 2021 年 7 月 12 日から 9 月 30 日のツイートに送られたリプライを収集した。Standard Search API [11]では、検索対象アカウントを指定しない形でのキーワードによる収集が可能だが、1 週間以内のツイートしか収集できないという制限がある。他方、Get Tweet timelines では、指定したアカウントの 3,200 件までのツイートが収集可能であり、同件数以内であれば期間についての制限はない。以上を踏まえ、データ収集時に 2021 年 7 月 5 日から一か月以上が経過していたため、次のような形で段階的にツイートを収集した。

はじめに、「医師」「専門家」「知事」の 9 月 1 日から 9 月 30 日のツイートに対するリプライを Twitter Standard API の Standard Search API を用いて複数回に分けて収集した。各アカウントのスクリーンネームをキーワードに用いて検索し、抽出されたツイートのうちリプライに対応するツイート ID とツイートを行ったアカウントのスクリーンネームを取得した。取得したスクリーンネームを用いて該当アカウントのツイートを取得し、対応するリプライを抽出した。

9 月 1 日以前のツイートに対するリプライについては、Standard Search API では 1 週間以上前のツイートは収集できないため、9 月 1 日から 9 月 30 日の「医師」「専門家」「知事」のツイートに対するリプライをしたアカウントのものに限定した。9 月 1 日から 9 月 30 日のリプライを行ったユーザーのスクリーンネームを用いて、Get Tweet timelines によりツイートを収集し、その中から 9 月 1 日以前のツイートに対するリプライを抽出した。収集されたリプライの総数は医師が 45,003 件、専門家が 43,976 件、知事が 98,544 件となった。

表1 医師, 専門家, 知事へ送られたリプライのスコアの平均・標準偏差・件数

	医師	専門家	知事
平均	-2.82	-2.99	-5.08
標準偏差	3.47	3.76	4.68
件数	45,003	43,976	98,544

4.3 感情値の比較

「医師」「専門家」「知事」のメッセージを受けたユーザーの感情を分析するために、そのリプライの感情値を算出する。ここでは、形態素解析エンジン MeCab[12]と高村ら[13]が公開している単語感情極性対応表を用いた。単語感情極性対応表は単語の極性値が記載されたものであり、極性値が高い単語ほどポジティブな単語、低いほどネガティブな単語とされる。これを用い、各リプライに含まれる単語の極性値を足し合わせたものを、リプライに対するスコアとした。スコアが高いほど肯定的なリプライ、スコアが低いものほど否定的なリプライとここでは解釈する。リプライにスコアを割り当てるにあたり、Python のライブラリである正規表現モジュール `re` と、絵文字を扱うことができる `emoji` を用いて不要な本文に関係のない情報の削除などの前処理を施した。具体的には、リプライに含まれる URL やハッシュタグ、@ (アットマーク)、絵文字を削除した。「医師」「専門家」「知事」のツイートへのリプライの感情値の平均、標準偏差、リプライ件数を表1に示す。

以下、リプライのスコアの平均を感情スコアと呼ぶ。感情スコアは「医師」が-2.82(3.47), 「専門家」が-2.99(3.76), 「知事」が-5.08(4.68)となっている。この順でユーザーがツイートに対してよりポジティブな単語を発しており、好意的な反応を示した可能性がある。この順序は奥原ら[2]の実験研究の結果と一致しているものの、「医師」と「専門家」の間の差は僅かとなっている。

5. BERT による行動の分析

本章では、BERT を用いて「医師」「専門家」「知事」による外出の自粛を求めるツイートを受けたユーザーが、実際に外出を自粛したかどうかを分析した結果を報告する。具体的には、BERT を用いて「医師」「専門家」「知事」に対してリプライを送ったユーザーの2022年7月5日から9月30日の全ツイートを「外出したことを示すツイート」と「その他のツイート」に分類し、外出したことを示すツイートを行ったユーザー数の割合と外出ツイート数の割合を非常事態宣言の前後等の期間で比較した。

5.1 識別モデル

本研究では、ニューラル言語モデルの一種である BERT を用いる。BERT は双方向の Transformer により、大規模なデータセットでモデルの事前学習を行い、少量の教師データを用いてファインチ

ューニングを行うことで、様々な自然言語処理のタスクにおいて高い性能を持つモデルが構築できることが知られている。ここでは、「外出したことを示すツイート」と「その他のツイート」を識別分類する2値識別モデルを、乾ら[14]による日本語 Wikipedia で事前学習を行った BERT モデルをファインチューニングすることにより構成する。実装には Python と深層学習のフレームワークである PyTorch [15]および自然言語処理などタスクにおける様々なアーキテクチャを提供する Transformers [16]を用いた。

5.2 「外出」に近い単語の抽出

収集したツイート群において、「外出」に近い意味で用いられる単語を特定し、その単語が含まれるツイートを抽出し、その中でも明らかに外出していたことを示すツイートを選別する方針をとった。はじめに、自然言語処理に関するライブラリ Gensim [17]を用いて、収集したリプライを学習データとした Word2Vec モデルを作成し、各単語の分散表現を生成した。データセット内において「外出」に近い意味を持つ単語の分散表現ほど、「外出」の分散表現とのコサイン類似度が小さくなることが期待される。Word2Vec の学習を行う際には、ハイパーパラメータの調整が必要となる。Mikolov ら[18]において、データセット内の単語数およびベクトルの次元の関係が調べられている。特に、学習データでの出現回数が5回以上の692,000単語に対しては、分散表現の次元を300と設定していた。また、[19]ではデータの量に応じ、ベクトルの次元を線形にスケールリングを行うことで性能が改善している。ここで用いる学習データでは出現回数が5回以上の単語は17,955単語であり、692,000単語の約39分の1の単語数である。これらを踏まえ、次元数を8、その他のパラメータは Gensim のデフォルト値を用いて Word2Vec の学習を行った。モデルを構築後、コサイン類似により「外出」に近い上位50件の単語を抽出した。「運動会」や「修学旅行」、「アルコール」、「行事」などの単語が抽出された。

5.3 識別モデルデータセットの構築

BERT によるファインチューニングのためのデータセットを次のように作成した。はじめに、2021年9月1日から9月30日の間に「医師」「専門家」「知事」にリプライを送ったことのある36,812ユーザーの2021年7月5日から9月30日の全ツイートを収集した。収集したツイートから5.2節の「外出」と近い意味で用いられる単語を一つ以上含むツイートを抽出した。

抽出したツイートから「医師」「専門家」「知事」ごとに20,000件ずつランダムサンプリングを行い、計60,000件のツイートをランダムに並べ替えた後、明確に外出したことを示すと判断できたツイートを先頭から250件選び、負例のラベルを付与した。また、外出と近い意味で用いられる単語が含まれるが、外出をしたことを示していないと判断できるツイートについても同様に250件選んでおき、正例のラベルを付与した。さらに、「外出」と近い意味で用いられる単語が含まれていなかったツイートについても250件を集め、正例として追加した。

5.4 識別のモデルの構築

作成したデータセットのうち、その6割を学習データとし、残り4割のうち2割ずつをそれぞれ

検証データとテストデータに割り当てた。学習データと検証データを用いて、乾ら[14]による日本語 Wikipedia で事前学習を行った BERT をファインチューニングすることでモデルを構築し、テストデータ上でモデルの評価を行った。ファインチューニングにあたっては、epoch 数は 10 とし、学習率を 0.00001 に設定した。

ファインチューニングにより構築したモデルの、テストデータにおける正解率および精度、再現率、F 値のマクロ平均を表 2 に示す。ランダムな識別器の正解率 50%と比較すると 35%向上しており、トレードオフの関係にある適合率と再現率の差は 1%となっている。

5.5 構築したモデルによる識別

前節で構築した識別モデルを用いて「医師」「専門家」「知事」へ 1 件以上リプライを送ったことがあるユーザーの 2021 年 7 月 5 日から 7 月 18 日までのツイートを「外出したことを示すツイート」と「その他のツイート」に分類した。東京都に 7 月 12 日に緊急事態宣言が出されており、その直後に知事、専門家、医師からの外出自粛を促すメッセージが多く発信されている可能性が高いと考えたためである。外出したことを示すツイートを行ったユーザー割合を算出したところ、宣言前後 2 週間のそれぞれのユーザー群の平均は、「医師」(6.1%)「専門家」(4.7%)「知事」(4.9%)であった。

5.6 外出ツイートの割合の時系列変化

はじめに、日ごとの時系列推移について観察した。初日の 7 月 5 日を基準値 1 とした場合の日別の変化を図 1 に示す。緊急事態宣言が予告された 7 月 8 日にいずれのユーザー群もいったん最低を記録しているが、徐々に増加し、再び宣言が実際に発令された 7 月 12 日に減少に転じている。

表 2 医師、専門家、知事へ送られたリプライのスコアの平均・標準偏差・件数

適合率	再現率	F 値	正解率
0.83	0.84	0.84	0.85

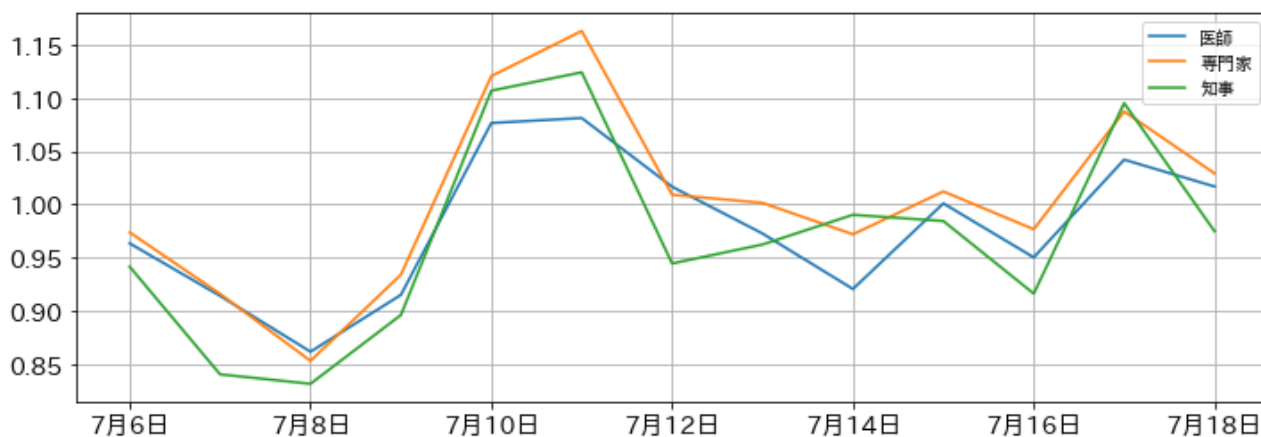


図 1 初日の 7 月 5 日を基準値 1 とした場合の外出ツイート割合の日別の推移

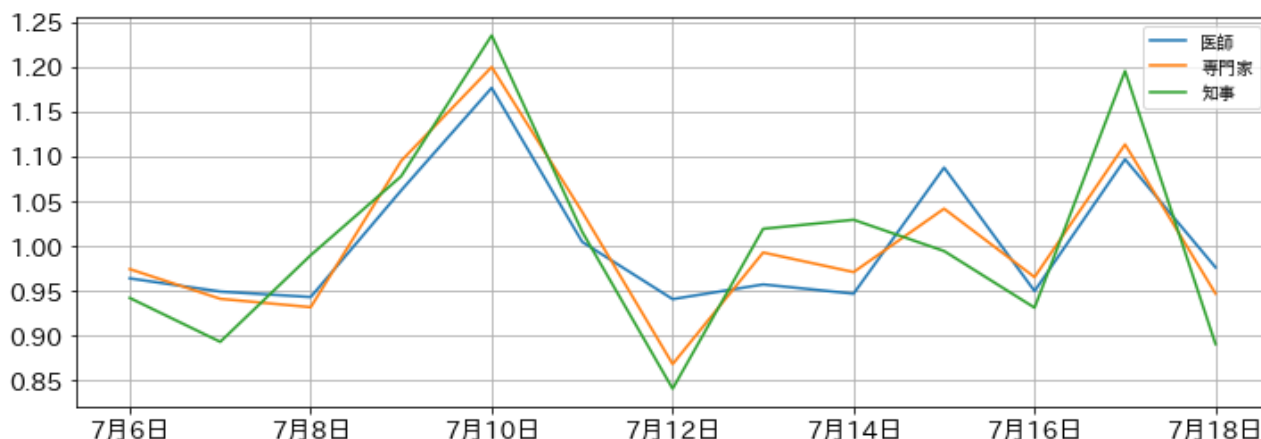


図2 外出ツイート数の前日比の推移

表3 7/5-7/11 から 7/12-7/18 における外出ユーザー数および外出ツイート数の増加率 (%)

	外出ユーザー の増加率	外出ツイート 数の増加率
医師	+0.43	+1.77
専門家	+0.45	+2.32
知事	+0.61	+2.43

なお、7月10日、11日および7月17日、18日はそれぞれ土曜日、日曜日となっていることにも注意されたい。

同期間における前日比の変化を図2に示す。前日比はこの日にいずれのユーザー群も最大値を記録しており、それぞれ「医師」(6.0%減)「専門家」(13.2%減)「知事」(16.0%減)であった。三者のユーザー群の7月5日時点に対する外出割合の順序は日によって異なる。

5.7 前後一週間における外出ユーザー数と外出ツイート数の割合比較

続いて、一週間単位での外出したユーザーの割合と外出ツイート数の割合を比較した。外出したことを示すツイートを行ったユーザー割合の宣言前後一週の変化(%)と外出ツイート数の割合の変化を表3に示す。

東京都に緊急事態宣言が発令された2021年7月12日の前後一週間では、外出したことを示すツイートはいずれも増加はしているものの、ユーザーの増加量の観点からは「医師」が一番低く、続いて「専門家」、「知事」の順となっており、奥原ら[2]と同じ順序で外出の増加が抑制されている。しかしながら、4の感情分析の結果と同様に、「医師」と「専門家」の間の差はやはり「知事」との差と比較するとやはり小さいものとなっている。外出したことを示すツイートについても、いずれも増加はしているものの、増加量の観点からはやはり「医師」が一番低く、続いて「専門家」、「知事」の順となっている。ここでは、「専門家」と「知事」の差が僅かとなっている。

6. 考察

奥原らの実験研究[2]で測定した「外出自粛の気持ち」の度合いは、行動的側面ではあるが、あくまでも態度としての評価的反応である。しかし、本研究の「行動」分析結果から、態度だけではなく実際の行動変容から、やはり「医師」「専門家」「知事」の順で説得効果が高かった可能性が示唆された。

このメッセージの送り手の順番に貢献している説得的コミュニケーションの要素として、1つは「信頼性」が考えられる。先行研究でも「医師」次いで「科学者」すなわち本研究における「専門家」の信頼性が高く、政治家すなわち本研究における「知事」は比較的低いという指摘とも合致する。もう1つは、本研究の「感情」の分析結果において同じ順番でメッセージに対しポジティブな単語が発せられていることから、「好感度」が関わっている可能性がある。特に、本研究ではメッセージの受け手がツイートのメッセージを読んだうえでリプライも行っていることから、もともとTwitterのフォロワーあるか、あるいは検索結果から選択的に読んでいることが推測され、受け手が送り手に好意をもっている可能性がある。

本研究の限界として、メッセージの送り手の「誰が」のみを検討しており「何を」伝えるか分析していない点は留意が必要である。外出自粛を促す内容は、非常事態宣言が発出された時期に発信されたツイートには含まれている可能性が高いが、実際にはすべては確認できていない。特に、知事はコロナ以外にも様々な行政に関するツイートを行うため、伝える内容の分散も大きいと考えられるが、ここではそれらも含めてユーザーの平均的な傾向を観察した。奥原らが実験で意図的に用いたメッセージに込めた、動物的に反応する心に刺さるメッセージなど行動変容に貢献する内容の検討も行っていない。したがって、どのような内容に対する「感情」であるか、「行動」であるかは不明で、因果関係については定かではないが、少なくとも相関は見出されている。

また、本研究の分析方法の限界から、次のような誤差が生じうることに注意されたい。まず、ファインチューニングを行う際に、確実に外出したと判断できたもののみを負例とし、それ以外のものはすべて正例としていた。たとえば、外出の予定をツイートしていても、外出中のツイートがないため外出していないとみなされたが、実際には外出していたアカウント等は存在しうる。また、感情値の際には単語感情極性対応表を用いた。文脈については考慮されていないため、実際はポジティブなツイートの感情値が低くでるということが起こりうる。ここでは、そのようなものも含めその平均を観察している。また、Twitterユーザー自体にも偏りが指摘されている[20]。今回の検証では「医師」「専門家」「知事」から確実にメッセージを受け取ったユーザーを対象とするために、リプライを行ったことがあるアカウントに限定している点にも留意が必要である。本研究で用いたデータの量は奥原ら[2]と比較し多いものの、あくまでTwitterユーザーの傾向に関する分析である。

7. 結論

日本語 Wikipedia で事前学習を行った BERT モデルをファインチューニングすることにより、外出したことを示すツイートの識別器を構築した。識別器を用い、東京都における緊急事態宣言の前

後において、「知事」「専門家」「医師」に対しリプライをしたことのあるユーザーごとに、外出したことを示すツイートの増加量を比較した。増加量は「医師」「専門家」「知事」の順に抑制されていた。この結果は奥原ら[2]による「外出自粛の気持ち」の度合いに関する結果と一致していた。また、単語感情極性対応表を用いて、緊急事態宣言の発令から解除までの期間における「知事」「専門家」「医師」に対するリプライの感情スコアを比較したところ、「医師」「専門家」「知事」の順に高い結果となった。

本研究の限界として、メッセージの送り手が「誰が」のみを検討しており、「何を」伝えるか分析していない点が挙げられる。行動変容の予測因子については、さらなる分析が必要である。技術的な観点からは、識別器のためのデータセットの構成から外出の判定が慎重になりすぎている可能性や、感情値の計算に文脈が考慮されていないこと等が挙げられる。

参考文献

- 1) S.V. Berg, J.K. O'Hara, M.T. Shortt, H.Thune, K.K. Brønnick, D.A. Lungu, J. Røislien, and S. Wiig, "Health authorities' health risk communication with the public during pandemics: a rapid scoping review," BMC Public Health, vol.21, no.1, pp.1401, July 2021. DOI: 10.1186/s12889-021-11468-3
- 2) T. Okuhara, H. Okada, and T. Kiuchi. "Examining persuasive message type to encourage staying at home during the COVID-19 pandemic and social lockdown: a randomized controlled study in Japan," Patient Education and Counseling, vol.103, no.12, pp. 2588-2593, Dec. 2020. DOI: 10.1016/j.pec.2020.08.016
- 3) 石川ひろの, "行動変容を促すコミュニケーション, " 保健医療専門職のためのヘルスコミュニケーション学入門, pp.55-66, 大修館書店, 東京, 2020.
- 4) M. Ricou, T. Pereira, H.P. Pereira, R. Picoli, and S. Marina. "Covid-19 pandemic: effect on confidence levels of portuguese towards people of different professions," Risk Management and Healthcare Policy, vol.2021, no.14, pp.4141-4148, Oct. 2021. DOI: 10.2147/RMHP.S310608
- 5) 奥原剛, 岡田宏子, 木内貴弘, "感染症とワクチンのコミュニケーションで「何を」伝えるか, " 産婦人科の実際, vol.70, no.3, pp.263-269, March 2021.
- 6) 鳥海不二夫, 榊剛史, 吉田光男, "ソーシャルメディアを用いた新型コロナ禍における感情変化の分析, " 人口知能誌, vol.35, no.4, pp.F-K45_1-7, July 2020.
- 7) 井原史渡, 岸本大輝, 栗原聡, "新型コロナウイルスに伴う Twitter の分析と感染状況との関連性, " 人工知能学会全国大会論文集, vol.2021, pp.114GS4c01, June, 2021.
- 8) 久光祥平, 豊田正史, 吉永直樹, 張翔, "Twitter ユーザの COVID-19 ワクチン接種に対するスタンスおよびその分極化の推移に関する分析, " DEIM 2022 (第 14 回データ工学と情報マネジメントに関するフォーラム) 最終論文集, <https://proceedings-of-deim.github.io/DEIM2022/>, pp.A24-6, Feb.2022.
- 9) J. Devlin, M.W. Chang, K. Lee, and K. Toutanova. "BERT: pre-training of deep bidirectional transformers for language understanding," Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, vo.1 (Long and Short Papers), pp. 4171-4186, Minneapolis, MN, June 2019.
- 10) Twitter. "Get Tweet timelines". https://developer.twitter.com/en/docs/twitter-api/v1/tweets/timelines/api-reference/get-statuses-user_timeline, 参照 Jan.24,2022.

- 11) Twitter. "Search Tweets: Standard v1.1". <https://developer.twitter.com/en/docs/twitter-api/v1/tweets/search/api-reference/get-search-tweets>, 参照 Jan.24,2022.
- 12) T. Kudo. "MeCab: Yet Another Part-of-Speech and Morphological Analyzer". <https://taku910.github.io/mecab/>, 参照 Feb.4,2022.
- 13) H. Takamura, T. Inui, and M. Okumura. "Extracting semantic orientations of words using spin model," Proceedings of the 43rd Annual Meeting of the Association for Computational Linguistics (ACL'05), pp.133-140, Ann Arbor, MI, June 2005.
- 14) Inui Laboratory. "cl-tohoku/bert-japanese". <https://github.com/cl-tohoku/bert-japanese>, 参照 Feb.22,2022.
- 15) A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga, A. Desmaison, A. Köpf, E. Yang, Z. DeVito, M. Raison, A. Tejani, S. Chilamkurthy, B. Steiner, L. Fang, J. Bai, and S. Chintala. "Pytorch: An imperative style, high-performance deep learning library," Advances in neural information processing systems, 32, pp. 8024-8035, 2019.
- 16) T. Wolf, L. Debut, V. Sanh, J. Chaumond, C. Delangue, A. Moi, P. Cistac, T. Rault, R. Louf, M. Funtowicz, J. Davison, S. Shleifer, P.V. Platen, C. Ma, Y. Jernite, J. Plu, C. Xu, T.L. Scao, S. Gugger, M. Drame, Q. Lhoest and A. Rush. "Transformers: State-of-the-art natural language processing," Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing, pp. 38-45, Online, Oct. 2020.
- 17) R. Řehůřek and P. Sojka. "Software framework for topic modelling with large corpora," Proceedings of the LREC 2010 Workshop on New Challenges for NLP Frameworks, Valletta, Malta, pp.45-50, May 22, 2010.
- 18) T. Mikolov, I. Sutskever, K. Chen, G. Corrado, and J. Dean. "Distributed representations of words and phrases and their compositionality," Advances in Neural Information Processing Systems, 26, pp. 3111-3119, Dec.2013.
- 19) T. Mikolov, K. Chen, G. Corrado, and J. Dean. "Efficient estimation of word representations in vector space," arXiv preprint arXiv:1301.3781, 2013. DOI:10.48550/arXiv.1301.3781
- 20) A. Mislove, S. Lehmann, Y.Y.Ahn, J.P. Onnela and J. Rosenquist. Understanding the Demographics of Twitter Users. Proceedings of the Fifth International AAAI Conference on Web and Social Media, 5(1), pp. 554-557, Barcelona, Spain, July 2021.

要 旨

本研究では、新型コロナウイルス感染症の拡大に伴う外出自粛期間の「医師」「専門家」「知事」からのツイートへのリプライに見るユーザーの「感情」および「行動」について、説得的ヘルスコミュニケーションの観点から分析した。分析には、三者にリプライしたことがあるユーザーのツイートと、単語感情極性対応表およびニューラル言語モデル BERT を用いた日本語版 Wikipedia により事前学習済みの BERT をファインチューニングすることで、ツイートが外出を示すものか否かを識別するモデルを構築した。構築した識別モデルにより非常事態宣言の前後における外出の増加量を調べた結果、「医師」「専門家」「知事」の順でメッセージの説得効果が高かった可能性が示唆された。また、単語感情極性対応表による感情分析の結果とも相関関係が観察された。

キーワード : BERT, Twitter, COVID-19, ヘルスコミュニケーション, 健康行動