

観光地に関するソーシャルメディア投稿の 分類手法の提案

櫛部 健汰 伊藤 恵

近年, ソーシャルメディアを観光時に活用することが増加している。観光地に関する情報を容易に収集できたり, 簡単にユーザーと共有できるなど近年の観光行動に必須になりつつある。しかし, ソーシャルメディア情報は信憑性が低く, 情報量が膨大であるというデメリットがある。観光従事者はデメリットの影響でソーシャルメディアから観光地に関する有益な情報を取得するのが難しくなっている。そこで本研究の目的として特定の観光地に関するソーシャルメディア情報をその観光地の観光従事者が容易に利用できるように, ソーシャルメディア情報分類手法を提案する。本研究では日本で利用者が多い Twitter に着目した。アカウントからキーワードを抽出し, そのキーワードをもとに Twitter 情報 (以下, Tweet) を収集する。収集した Tweet に関してトピックモデルを活用することでトピックごとに分類したのちに観光従事者が活用しやすいように動機 Tweet, 移動願望 Tweet, 現地 Tweet, 評価 Tweet の 4 種類に分類する。

In recent years, the use of social media has been increasing, and it has become indispensable for tourism activities, as it allows us to easily collect information about tourist attractions and share it with other users. However, social media information has some disadvantages such as low credibility and huge amount of information. The disadvantages make it difficult for tourist workers to obtain useful information about tourist attractions from social media. Therefore, the purpose of this study is to propose a social media information classification method to make it easier for tourist workers to use social media information about a specific tourist attraction. In this study, we focused on Twitter, which has many users in Japan. We extract keywords from accounts and collect Twitter information (hereinafter referred to as "Tweet") based on the keywords. The collected Tweets are categorized by topic using a topic model, and then classified into four types: promotional Tweet for tourist attraction, wishing to visit or moving to visit Tweet for tourist attraction, local Tweet of tourist attraction and evaluation Tweet of tourist attractions.

1 背景

近年, スマートフォンの普及に伴い, 個人から企業まで様々な情報を受け取り, さらには発信することができる情報社会となっている。令和元年度版情報通信白書[10]によると, 2018 年度において日本人のスマートフォンの保有率は約 8 割を超えている。インターネットに接続する情報機器としてはパソコンをすでに上回っている。スマートフォンの利用拡大の原因

の 1 つとしてソーシャルメディアの普及がある。

ソーシャルメディアとは, インターネットにて誰もが参加可能で社会的相互関係を生み出し利用できるメディアである。従来の TV や新聞などのマスメディアとは異なり, 誰もが情報発信ができるのが特徴である。代表的なソーシャルメディアとして Facebook, Twitter が挙げられる。日本では, ソーシャルメディアの利用者は年々拡大しており, 総務省情報通信制作研究所[11]の調査によるとスマートフォンを所有している人のうち Facebook を活用している人が 32.7%, Twitter を活用している人が 38.7% となっていた。さらに, 観光中にソーシャルメディアを活用することも増加している。株式会社 JTB 総合研究所[5]の調査によると直近の旅行でソーシャルメディア投稿を行なった人は 26%, 特に 29 歳以下の人に限定すると男性

A Proposal for a Classification Method of Social Media Posts about Tourist Attractions.

Kenta Kushinobe, 公立はこだて未来大学大学院 システム情報科学研究科, Future University of Hakodate, Graduate School of System Information Sciences.

Kei Ito, 公立はこだて未来大学, Future University of Hakodate.

が 42.4 %，女性が 47.6 %となっている。観光従事者は観光地に対する意見や評価をソーシャルメディアから得ることができればリアルタイムに評価を得ることができる。観光地に関するソーシャルメディア情報を分析することで、観光客へ直接アンケート等で調査を行うよりコスト等が少なくしつつ観光地の意見を得ることができると期待されている。

しかし、ソーシャルメディアの情報に関する問題点として信憑性が低く、さらに情報量が多いといった点がある。観光従事者が観光に関するソーシャルメディア情報を活用しようとすると、信憑性が低く膨大な情報量の中から有用性のある情報を観光従事者が効率良く取得することは難しい。

前述した課題や背景から、本研究では観光地に関するソーシャルメディア情報を効率的に取得できるような手法を提案する。アプローチとして本研究ではソーシャルメディアの一つである Twitter に着目し、検索ワードから Twitter 情報（以下、Tweet）を取得する。取得した Tweet を観光地に活用しやすいようにトピックごとに分類し、トピック分類した Tweet を観光従事者が活用しやすいように分類を試みる。分類した Tweet を観光従事者に提供することで直接 Twitter から取得するよりも容易に取得でき、活動の補助ができるのではないかと考えた。

本稿では第 2 章で関連研究について述べる。第 3 章では提案システムの概要について述べ、第 4 章で適用例並びに今後の使用実験について述べる。最後に第 5 章では本稿のまとめ並びに今後の展望について述べる。

2 関連研究

本章では関連研究について述べる。関連研究としてインターネット上から観光情報を抽出する研究、ソーシャルメディア情報を分類する研究などが挙げられる。

2.1 観光情報抽出に関する研究

近年、観光地に関する情報を口コミやソーシャルメディアから効率的に抽出を試みる研究は多く行われている。

渡邊らの研究 [13] の研究では観光地名が入っていな

い Tweet から観光地に関する感想の自動抽出手法を提案している。観光地名を含んでいる Tweet を教師データとして活用し観光地名を含んでいない Tweet に関しても感想を抽出できることがわかった。小原ら [8] の研究では Tweet 本文を用いて観光情報の抽出手法を提案した。観光情報を抽出するために日本語のパターンマッチングと観光に関する Tweet で用いられやすい表現などから観光地に関する Tweet の抽出を試みた。どちらの研究とも、Twitter から観光情報を抽出する点は類似しているが、本研究では取得だけではなく観光従事者に活用しやすいように分類も試みる。

2.2 ソーシャルメディア情報の分類に関する研究

ソーシャルメディア情報を活用しやすいように分類する研究も多く行われている。

橋田ら [4] の研究では観光地の改善点や PR 点を分析するために Twitter 上に投稿される観光地に関する Tweet に着目し畳み込みニューラルネットワークを用いて分類する手法を提案した。観光地に関する Tweet を観光地に意見しているかしていないかに分類を行なっている。本研究と観光地に関する Tweet を分類することは類似しているが、教師データを用意することが必要となっている。本研究では教師データを必要としない分類手法を提案する。

3 提案システム

本章では本研究の提案システムについて述べる。図 1 に提案システムの概要図を示す。提案システムの目的として観光地に関する Tweet（以下、観光 Tweet）を収集し、観光従事者が活用しやすいように分類し提供することとする。提案システムは 5 つのステップに分ける。ステップは以下の通りである。

- ステップ 1 観光地サイトからの検索ワード抽出
観光地サイトから観光地に関する検索ワードを抽出する
- ステップ 2 Tweet 収集
ステップ 1 で抽出した検索ワードを用いて Twitter Search API を用いて Tweet を収集する
- ステップ 3 トピック分類

ステップ2で収集した Tweet をトピックごとに分類し、ステップ4で選択したトピックワードに当てはまる Tweet を提示対象とする。

ステップ4 トピックワードの決定

観光従事者のアカウントから観光地に関するトピックワード候補を観光従事者に提供し、観光従事者にトピックワードを選択してもらう

ステップ5 観光 Tweet 分類

トピック分類した Tweet を有用に使えるように動機 Tweet, 移動願望 Tweet, 現地 Tweet, 評価 Tweet に分類する。

次節からは各ステップの詳細を説明する。

3.1 観光地サイトからの検索ワード抽出

まず、Twitter で効率良く観光 Tweet を収集するために提案システムでは観光地サイトから観光地に関するキーワード(以下、検索ワード)の抽出を行なった。単に観光地名だけで検索し Tweet を収集しても観光地に関係しない数多くの Tweet も収集してしまう。そこで検索ワードを効率良く収集するために検索ワードの類型化を定義した。表1に類型化した検索ワードとその例を示す。「観光地の特定用語」とは特定の観光地の名産品、施設名、グルメなどを指す。函館を例に挙げると「ラッキーピエロ」、「五稜郭公園」などを指す。「観光地名+観光用語」とは観光地名と観光行動に即した固有名詞、名詞を指す。函館を例に挙げると「函館 旅行」、「函館 観光」などを指す。「観光地名+観光地特徴語」とは観光地名と前述の「観光地名+観光用語」では抽出できない特定の観光地で使われることが多い名詞を指す。函館を例に挙げると「函館 海鮮」、「函館 イカ」などが挙げられる。類型化した検索ワードを用いて Tweet を収集する。

3.2 Tweet 収集

3.1節で取得した検索ワードを使用して Twitter Search API を用いて Tweet の収集を行う。収集した Tweet はデータベースに保存しておき、ステップ3で使用できるように保存しておく。収集した Tweet に関しては次ステップで活用しやすくするために、以下の制約を設けて前処理を行なった。

表1 検索ワードの類型化と活用例

検索ワードの類型	活用例
観光地の特定用語	ラッキーピエロ 五稜郭公園
観光地名+観光用語	函館 旅行 函館 観光
観光地名+観光地特徴語	函館 海鮮 函館 イカ

1. 語彙の品詞は、固有名詞・一般名詞・サ変接続の名詞のみ考慮する
2. 語彙は、2字以上漢字・ひらがな・カタカナ・数字の組み合わせで構成されている
3. リツイートを示す「RT」・リプライを示す「@username」・ハッシュタグを示す「#」に続く文字列・URLを除く

3.3 トピック分類

収集した観光 Tweet はトピックごとに分類を行う。分類を行うことで3.4節にて後述するトピックワードのパターンマッチングだけよりも幅広い Tweet を対象とすることが期待できる。観光 Tweet のトピックごとの分類を行うために本研究ではトピックモデルを用いる。トピックモデルとは文書に含まれる潜在的な話題を確率的な生成モデルで表現し、この生成モデルの推論を行う手法である。トピックモデルの代表的な手法の一つとして Blei ら [1] が提唱した LDA (Latent Dirichlet Allocation) がある。この手法では、一つの文章に複数の潜在的トピックが存在することを仮定し、各トピックおよびトピックで発生する単語を多項分布でモデル化する。そしてこのモデルのパラメータを推論アルゴリズムによって推定し、それを文書データに潜在的に存在する話題の分布とそれぞれの話題の単語の分布とする。しかし LDA は Tweet のような1文章に含まれる単語数が少ない文章には学習が困難であることが指摘されている [12]。そこで提案システムでは BTM (Biterm Topic Model) [2] を用いて Tweet の話題分類を試みる。BTM とは文書を biterm と呼ばれる非順序の二単語対に変換し、各 biterm のトピックを推論することによって文書に含まれるトピッ

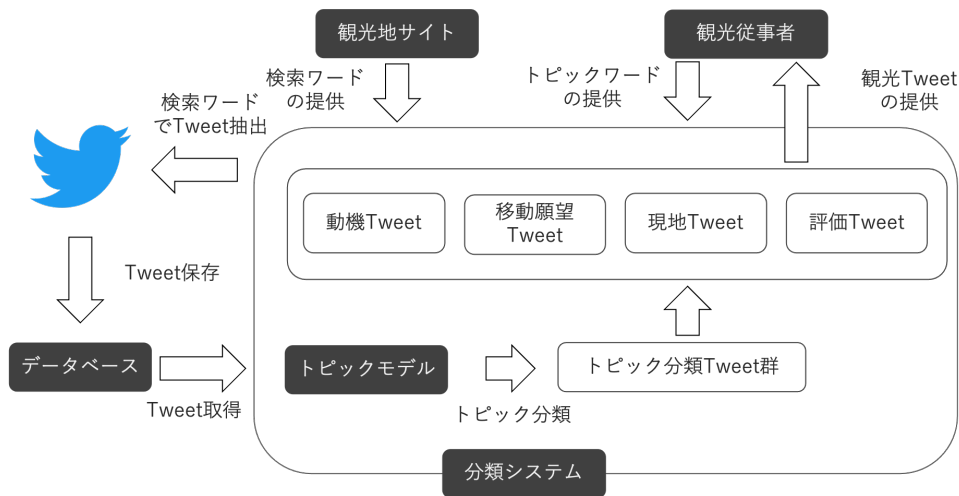


図1 提案システムの概要図

クを学習する手法である。Cheng らは、BTM によって推定されるトピックが、既存のトピックモデルと比較して、トピックに出現する語の一貫性がより高いことを報告している [2]。図2 に文書から抽出される Biterm の例を示す。提案システムでは BTM を用いて観光 Tweet のトピック分類を試みる。

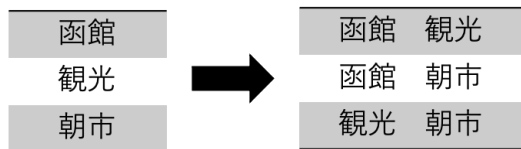


図2 Biterm の例

3.4 トピックワードの決定

収集した観光 Tweet を観光従事者が活用するためにその観光従事者の特徴を示すキーワード (以下、トピックワード) の指定が必要となる。提案システムでは観光従事者の Twitter アカウントからトピックワードの候補を提示する。観光従事者の Twitter の自己紹介文と過去 60 日以内の Tweet を活用し、それぞれ形態素解析を行った後に TF-IDF 値を算出して観光従事者にトピックワードの候補を提供する。提示さ

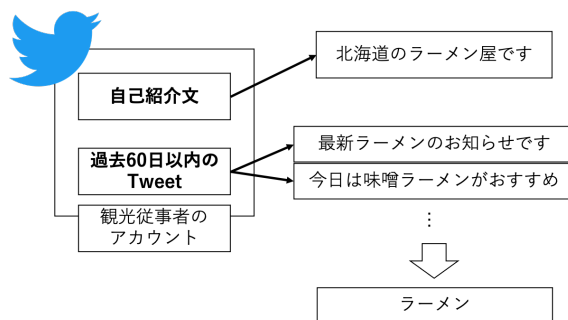


図3 トピックワード抽出イメージ図

れたトピックワードの候補を参考にしながら、観光従事者には自らトピックワードを選択してもらう。図3 にトピックワード抽出のイメージ図を示す。図3 では選んでもらったトピックワードに基づく Tweet を分類対象 Tweet とする。

3.5 観光 Tweet 分類

分類対象 Tweet に対して観光従事者が活用しやすいように観光 Tweet の分類を行う。分類にあたって金ら [6] の研究を参考に観光 Tweet の類型化を行なった。表2 に観光 Tweet の類型化とその例を示す。動機 Tweet, 出発願望 Tweet, 評価 Tweet, 現地 Tweet の4 種類に分類を行う。図4 に Tweet の分類順を示

表 2 観光 Tweet の類型化

観光 Tweet	説明	例
動機 Tweet	観光地へ動機付けを印象付ける Tweet	五稜郭公園では桜が見頃です
出発・願望 Tweet	観光地への願望, 観光地に移動中の段階の Tweet	五稜郭公園に行きたい
現地 Tweet	観光地の現地情報がある Tweet	五稜郭公園で桜が咲いている
評価 Tweet	観光地の評価情報がある Tweet	五稜郭公園の桜が綺麗だ

す。動機, 出発願望, 評価, 現地の順番に Tweet 分類し, 抽出する。抽出条件に満たさなかった Tweet に関しては観光従事者に提示外とする。

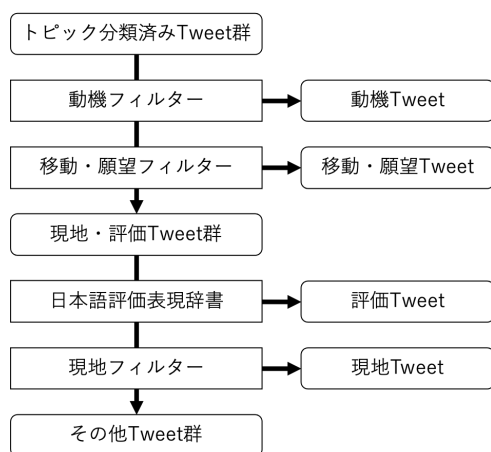


図 4 観光 Tweet の分類順

動機 Tweet とは観光地への動機をつける Tweet を指す。提案システムでは観光地の公式アカウントを事前に登録しておき, そのアカウントの Tweet は全て動機 Tweet として分類する。

出発・願望 Tweet とは観光地への願望を投稿, 観光地へ現在向かっている移動段階での投稿を指す。提案システムでは文末に着目する。文末に「行きたい」, 「～をやりたい」, 「向かっている」などの表現に着目し分類する。さらに「到着」, 「着いた」, 「出発」といった場所を起点とした行動の発言がある場合にも出発・願望 Tweet と定義する。

評価 Tweet とは観光地への評価表現がある Tweet を指す。本研究では小林ら [7] の評価値表現辞書を用いる。評価値表現辞書に当てはまる単語の例として「おいしい」といった形容詞, 「絶品」といった名詞な

どが挙げられる。評価値表現辞書に当てはまる単語がある場合には評価 Tweet として抽出する。

現地 Tweet とは観光地の現地情報がある Tweet を指す。提案システムでは鬼塚らの研究 [9] を参考に分類を試みる。鬼塚らの研究 [9] では時間帯によるフィルタリングと言語情報や Twitter 特有の情報を用いたフィルタリングを行っている。時間帯によるフィルタリングでは深夜帯の投稿を除外する (0~4 時)。観光客の深夜帯の投稿は雑談や 1 日を振り返る投稿であると考えられるので現地 Tweet ではないと判断する。Twitter 特有の情報に関しては「～なう」のような「今」を表す投稿, 「～にいる」のような現在の行動を投稿に関しては現地 Tweet として定義する。

4 システム適用例

本章では提案システムの適用例について説明する。対象地域として函館市を用いた。

4.1 活用例

検索キーワードの設定は函館市の公式観光サイトであるはこぶら [3] から抽出を行なった。はこぶらでは観光地の名所, 土産物, グルメなどの単語を手動で抽出した。さらに, 「函館 海鮮」のような観光地特徴語の抽出も手動で行った。「函館 旅行」のような観光用語に関しては現段階では「観光」の類語を調べて定義した。検索キーワードを指定後, Twitter Search API を用いて観光 Tweet の取得を行った。Tweet 取得期間は 2021/7/23~2021/7/30 の 1 週間に指定した。取得した Tweet の一部を図 5 に示す。結果として, 7082 件の Tweet を取得した。取得した Tweet は 3.2 節で述べた前処理を行いデータベース上に保存した。

次にトピックの分類を行った。収集した 7082 Tweet

表 3 北海道・北東北の縄文遺跡群の Tweet 例

Tweet 本文
祝・世界遺産登録!!「北海道・北東北の縄文遺跡群」 / 函館・南北海道観光ガイド
縄文遺跡群の世界文化遺産登録に沸く 北海道内 函館ではお得な宿泊プランも登場
速報 函館に世界遺産が誕生したぞす〜っ!!
大船遺跡見てきたけど、観光バス通るのも一苦勞な場所だし、かなり小規模。 道路の拡幅も急斜面などところなので望めない。しかも近隣にホテルがひろめ荘しかないという…。 函館市よ。課題は盛り沢山だぞ。とりあえずひろめ荘のことまでバイパス繋げきらんと。
史跡大船遺跡 定時解説 - 体験 函館市公式観光情報サイトはこぶら

表 4 ランチの Tweet 例

Tweet 本文
ラッキーピエロは素晴らしい...
ラッキーピエロ通いたい...
今日は有給休暇を取って道の駅めぐり 余市、赤井川、神恵内、 岩内、蘭越、寿都、島牧、黒松内、森の道の駅をまわってきました で、今は森町のラッキーピエロで遅めの昼食です
五島軒のカレーテイクアウトしたかったけど テイクアウトしてなさそうだったから泣く泣く諦めた
今日のお昼ご飯は!! 函館市川原町にあります 函館豪一めん川原店さんで 油そば大盛り 920 円 ご馳走さまでした 次は辛いやつにしようかな
本日の朝ごはん ハセストの期間限定メニューにしてみた やきとり弁当のカレーごはん

表 5 五稜郭の Tweet 例

Tweet 本文
五稜郭タワーたけー
五稜郭来たけどタワーから見ると900もするからやめた ちなみにこんな感じらしい! …うん、知ってるか笑#日本一周カブの旅#カブで日本一周#
函館五稜郭の中にある箱館奉行所。しずかで癒される公園です。
函館帰ってからスタバよく行く。五稜郭駅前店は初めて! ドライブスルーあります
函館カフェ函館2日目…松前に行こうと思ったら結構距離が あるらしいので断念 15 時には洞爺湖のホテルに チェックインなので函館市内観光に変更。とりあえず五稜郭行くかぁ

に対し BTM を用いて、15 トピックに分類を行った。分類したトピック分類の Tweet 例を表 3、表 4、表 5 に示す。2021 年に世界遺産に指定された北海道・北東北の縄文遺跡群についてのトピックや、函館のランチについてのトピック、五稜郭公園に関するトピックなどに分類された。

次に検索ワードのトピックワードの抽出を行った。活用例では函館のご当地グルメの一つであるラッキーピエロのアカウントを活用した。ラッキーピエロのアカウントから 3.4 節の手法を活用しトピックワードの抽出を行った。結果として、ラッキーピエロのアカウントから「ラッキーピエロ、函館」というト

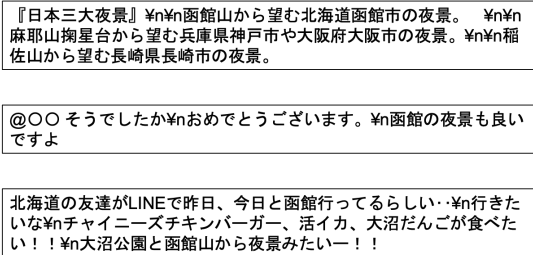


図 5 取得 Tweet 例

ピックワードを抽出した。トピックワードを用いて一番マッチするトピック分類済み Tweet を選択した。ラッキーピエロのトピックにマッチする Tweet 総数は 1130 件であった。

最後に観光 Tweet の分類を行った。まず、動機 Tweet の分類に当たって公式アカウントを登録する必要がある。公式アカウントは函館市の公式アカウントとラッキーピエロ公式アカウントを登録した。次に移動願望 Tweet の分類を行う。その後、評価 Tweet・現地 Tweet の分類を行った。各 Tweet の分類総数を表 6 に示す。トピック Tweet 総数 1130 件のうち動機 Tweet が 43 件、移動願望 Tweet が 467 件、評価 Tweet が 67 件、現地 Tweet が 34 件、その他 Tweet が 519 件であった。表 7、表 8、表 9、表 10 に各分類 Tweet の例を示す。

表 6 分類結果 Tweet の総数

観光 Tweet 類型	総数
Tweet 総数	1130
動機 Tweet	243
移動・願望 Tweet	467
評価 Tweet	67
現地 Tweet	34
その他 Tweet	319

4.2 提案システムの課題と今後

分類結果からトピックの分類は分類できているものと分類できていないものがあった。表 11 に分類が失敗しているトピックの例を示す。表 11 のトピックにはホテル関係と函館のお寿司屋さんの情報が混ざっている。トピック分類に関してはデータ量を増やすことでトピック分類の精度が上がると考えられる。トピック分類に関してはトピック数を 15 という値に設定して分類したが、15 が最適なトピック数とは限らないので最適なトピック数の設定も行っていきたい。観光情報分類にも、課題が見つかった。動機情報に関しては函館市の観光協会や函館地域の公式アカウント等を設定したが、アカウントの設定が曖昧であると動機情報が的確に定義できない。そこで観光従事者が公式アカウントを的確に設定できるようにアカウントガイドラインの策定が必要であると感じた。分類できなかった情報にも評価情報ではないかと思われる Tweet も散見された。例を挙げると「ハセガワストア!!! 函館の誇り!!!」といった Tweet があった。しかし、日本語評価辞書に当てはまる単語がないために評価情報と分類されなかった。今後評価情報の分類には単語以外にも着目の必要がある。現地情報の分類に関しては文末表現ではないが現地情報に含まれるべき Tweet を分類できていなかった。今後文末表現以外にも分類条件として追加する必要がある。現地 Tweet として分類できないものとして「観光施設名のみ+写真」という観光 Tweet が多く見つかった。そこで写真があるかの条件も現地 Tweet の分類条件として追加する。

今後取得した観光情報は観光従事者に提示する。提示手段として LINE Bot を用いて提示を行う。図 6 に提示イメージ図を示す。

LINE Bot では定期的に観光 Tweet を提示することで、観光従事者が観光地に関するトピックを定期的に取得することができる。取得することで Twitter から情報を直接取得するよりも容易に取得できるのではないかと期待している。LINE Bot 開発後は活用実験を行う。実際の観光従事者に活用していただき評価実験について述べる。提案システムの有用性を確かめるために使用実験を行う予定である。使用実験の評価指標としてアプリケーションの使用感、通知のタイ

表 7 動機 Tweet の例

Tweet 本文
函館にお越しの際には、ぜひ足を運んでみてください！ 公式 HP やホテルからのアクセスはこちらから。 ▼ラッキーピエロ（戸倉店は施設より車で 6 分） ▼プティ・メルヴィーユ（JR 函館駅前すぐ）
通販 #パーティー一度は食べてみたいチョコレート：radio（ラジオ）で有名 ★五島軒 コーンポタージュ（1 人前）
8/1 限定店内全品 5 % OFF！ 函館・五島軒の極上 ほぐし肉カレー 4 食セット 1 日 100 セット限定 送料無料 ネコポス便 お試し [楽天]

表 8 移動願望 Tweet の例

Tweet 本文
函館に行きたい！ ラッキーピエロに行きたい！
ラッキーピエロに行きてえ
ラッキーピエロ食べたい

表 9 評価 Tweet の例

Tweet 本文
昨日の試合のお昼は 函館五島軒さんの イギリス風ビーフカレー。 大会価格でお得に美味しく食べれました。 #モルック函館大会
ラッキーピエロ美味しかった
誕生日企画のお料理やりました！ ラッキーピエロのチャイニーズチキンバーガー！はちゃめちゃに美味かった 矢澤おめでとう！
ハセガワストアのやきとり弁当がまじでうまい。 ビール欲しかった

表 10 現地 Tweet の例

Tweet 本文
4 連休函館ラッキーピエロ総本店混んでます。 #函館#ラッピ#ラッキーピエロ#総本店#四連休 #ハンバーガー
函館のラッキーピエロ行ったよ！（味は覚えてない
ラッキーピエロ来たお
ラッキーピエロに着いたー

ミング、観光 Tweet の有用性を調査する。

5 まとめと今後の課題

本研究では観光従事者がソーシャルメディア情報を効率よく活用できるようにソーシャルメディア投稿を分類する手法を提案した。本研究ではソーシャルメディアの中でも Twitter に着目し、観光 Tweet を観

光従事者が活用しやすいように動機 Tweet、移動願望 Tweet、評価 Tweet、現地 Tweet に分類した。今後の展望としては大きく 3 つある。1 つ目はシステムの開発の継続である。現段階では Tweet の収集、分類トピック数の決定など手動で行っている部分が多くある。今後使用実験を行う上でも様々な部分の自動化を行っていく。2 つ目は実際の観光従事者への使

表 11 分類失敗トピックの Tweet 例

Tweet 本文
やっぱり函太郎の寿司は最高だな……(´ω´)
ラビスタ函館の朝食と言えば…コレコレ!イクラ掛けちゃうぞお#北の番屋
これ、セコマとラキビのメニュー一通り頼んでホテルで楽しむだけで旅行になるよね? めちゃうちゃニッチな函館オフする?しないよね知ってる

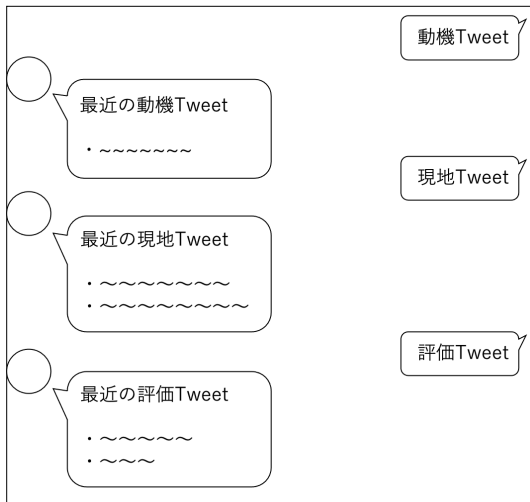


図 6 LINE Bot の提示イメージ図

用実験である。第 4 章でも述べたような実験を行い、システムの評価を行なっていく。3 つ目はソーシャルメディアの信憑性に関する検討である。ソーシャルメディアは多くの人が利用している反面、悪意のあるユーザーなどのデマ情報なども多く見受けられる。近年のコロナ禍ではデマ情報から観光地への風評被害も発生している。そこで提案システムに情報の信憑性も観光従事者に提供し、デマ情報の否定行動の手助けを行う。観光 Tweet の信憑性の向上を行うことで、観光地の風評被害の軽減を行うことを目指す。

参考文献

[1] Blei, D. M., Ng, A. Y., and Jordan, M. I.: Latent dirichlet allocation, *the Journal of machine Learning research*, Vol. 3(2003), pp. 993–1022.

[2] Cheng, X., Yan, X., Lan, Y., and Guo, J.: BTM: Topic Modeling over Short Texts, *IEEE Trans. Knowl. Data Eng.*, Vol. 26, No. 12(2014), pp. 2928–2941.

[3] 函館市観光部: 函館市公式観光情報サイトはこぶら, <https://www.hakobura.jp/>(最終アクセス 2021/08/09).

[4] 橋田修一, 田村慶一, 酒井達弘: 畳み込みニューラルネットワークを用いた観光ツイートの分類手法, *人工知能学会全国大会論文集*, Vol. JSAI2018(2018), pp. 1–4.

[5] 株式会社 JTB 総合研究所: コロナ禍の生活におけるインターネットや SNS からの情報, <https://www.tourism.jp/wp/wp-content/uploads/2019/11/smartphone-travel-consumption.pdf>, 2019(最終アクセス 2021/08/09).

[6] 金徳謙, 村越友香: 瀬戸内圏の地域文化の発見と観光資源の創造「直島の事例にみる SNS 上に流通する観光情報の類型化」, 美巧社, 2010.

[7] 小林のぞみ, 乾健太郎, 松本裕治, 立石健二, 福島俊一: 意見抽出のための評価表現の収集, *自然言語処理*, Vol. 12, No. 3(2005), pp. 203–222.

[8] 小原基季, 森田和宏, 泓田正雄, 青江順一: Twitter 本文を用いた観光情報抽出及び分析システムの構築, *人工知能学会全国大会論文集 第 29 回全国大会 (2015)*, 2015, pp. 1–3.

[9] 鬼塚友里絵, 嶋田和孝: 前後文脈を考慮した Tweet の現地性判断, *信学技報*, 社団法人 電子情報通信学会, (2014).

[10] 総務省: 令和元年度版情報通信白書, <https://www.soumu.go.jp/johotsusintokei/whitepaper/ja/r01/html/nd232110.html>, 2019(最終アクセス 2021.08.).

[11] 総務省情報通信制作研究所: 令和元年度情報通信メディアの利用時間と情報行動に関する調査報告書, https://www.soumu.go.jp/main_content/000708016.pdf, 2020(最終アクセス 2021/08/09).

[12] Tang, J., Meng, Z., Nguyen, X., Mei, Q., and Zhang, M.: Understanding the Limiting Factors of Topic Modeling via Posterior Contraction Analysis, *Proceedings of the 31st International Conference on Machine Learning*, Xing, E. P. and Jebara, T.(eds.), *Proceedings of Machine Learning Research*, Vol. 32, Beijing, China, PMLR, 2014, pp. 190–198.

[13] 渡邊小百合, 吉野孝: 観光地名なしツイートからの観光地に関する感想の抽出手法, *情報処理学会論文誌*, Vol. 59, No. 1(2018), pp. 43–51.