

# 半教師あり学習を用いた要望ツイートの抽出手法の評価

川島 崇秀<sup>1</sup> 佐藤 哲司<sup>1</sup> 神門 典子<sup>2</sup>

**概要:** 情報発信, 情報共有ツールとして定着している SNS の一つである Twitter 上では, 製品やサービスに関する口コミが日々大量に投稿されている. このため, Twitter 上の投稿を企業のマーケティング活動に活かそうという試みが注目されている. 本研究では, Twitter 上に投稿される口コミの中でも消費者のニーズを直接的に示す要望に注目し, Twitter から要望を含む投稿を抽出する手法を提案し, 実装・評価を行う. ツイートには文法的に崩れた表現や多様な文章表現が多く存在することから, 従来の辞書ベースの手法では抽出できる要望に限界があった. そこで, 要望を含む投稿の抽出に教師あり機械学習のアルゴリズムである SVM を適用するとともに, 学習データの収集に半教師あり学習の一つである Distant Supervision の適用を試みた. ソーシャルゲームに関する口コミ情報を対象として, 適合率・再現率・F 値を用いて従来手法との比較を行った結果, 提案手法の有効性が確認されたので報告する.

## Evaluation of Demand Tweet Extraction Method by using Distant Supervision

TAKAHIDE KAWASHIMA<sup>1</sup> TETSUJI SATOH<sup>1</sup> NORIKO KANDO<sup>2</sup>

### 1. はじめに

近年, ソーシャルメディアの普及により, 誰でも簡単に情報発信ができるようになった. ソーシャルメディア上では, 製品やサービスに関する口コミが日々大量に投稿されている. こういった背景から, 企業においてソーシャルメディア上の口コミ情報を市場調査や反響測定などのマーケティング活動に活かそうという試みが注目されている [1].

ソーシャルメディアの中でも近年顕著な普及を見せる Twitter は, リアルタイム性の高さ・ユーザの多様性・投稿量の多さから分析対象として大きな注目を集めている [2]. しかし, Twitter 上では, 日々何千万ものクチコミ情報が秒単位で行われており, これらの投稿を手手で分析することには膨大なコストが掛かる. したがって, Twitter 上の投稿をビジネス活動に活用するには, 膨大な量のツイートの中から, 有益なツイートを精度よく抽出する手法が不可欠である.

Twitter 上の投稿から有益なツイートを抽出する研究と

しては, センチメント分析が知られている [3]. センチメント分析とは, Twitter 上のレビューを肯定的あるいは否定的のいずれか2つのカテゴリに分類する手法である. しかし, 肯定的か否定的かに分類することは必ずしも, ビジネス活動に於いて価値のある情報を抽出することにはならない. 例えば, 要望を含む投稿は消費者のニーズを直接的に表す重要な情報であると考えられるが, センチメント分析の手法で抽出することはできない. 本研究では, 「要望」という点に着目して, 要望を含むツイートの抽出を試みる. 従来の手法で抽出の対象としていなかった要望を含むツイートを抽出することで, それらのデータをビジネス活動に活かしていくことが期待できる.

Twitter 上の投稿から商品やサービスに関する要望を含むツイートを抽出する際に課題となるのは, Twitter 上の投稿における多様な文章表現である. Twitter 上には文法的に崩れた表現や多様な文章表現が多く存在している. このため, 従来手法で提案されている辞書ベースの手法 [4] では, これらの多様な表現への対応が困難であった.

本研究では, 要望を含むツイートの抽出に機械学習のアルゴリズムを適用することで, 従来手法と比較してより高い精度での抽出を試みる. また, 学習データの収集に半教

<sup>1</sup> 筑波大学図書館情報メディア研究科 〒305-8550 茨城県つくば市春日 1-2

<sup>2</sup> 国立情報学研究所 情報社会相関研究系 〒101-8430 東京都千代田区一ツ橋 2-1-2

師あり学習の一つである Distant Supervision を適用することで、簡便に学習データを増やすを試みる。

## 2. 関連研究

Twitter 上の口コミ情報の抽出に関する研究は盛んに行われている。Twitter 上の投稿に対してセンチメント分析を行った研究としては、野畑ら [3] の研究が挙げられる。野畑らは教師あり機械学習の手法を用いて Twitter 上の投稿をポジティブなもの、ネガティブなもの、2つのカテゴリに分類している。

また、Twitter 上から要望を含む投稿を抽出する研究も行われている。栗原ら [4] は Twitter 上から地方自治体に関する要望を含む投稿の抽出を試みている。そこでの手法は、あらかじめ作成した要望表現の特徴を含む辞書を用いたパターンマッチングである。山本 [5] らは、Twitter を対象として生活に関連する単語からなる辞書を作成し、特定の地域の要望を含む、生活情報を抽出する手法を提案している。

要求表現の言語学的な定義に関する試みも行われている。大森 [6] は要求表現の定義として、「命令」、「依頼」、「禁止」、「誘いかけ」、「希望」、「当為」、「当为非断定」、「希望非断定」の態度を帯びる文は要求文であるとした。大森はさらに、要求表現の文法的な特徴として、「～しろ」「～たい」「～ほしい」といった文末表現を挙げている。

Web からの情報抽出のタスクに Distant Supervision を用いた研究も行われている。Distant Supervision [7] とは、半教師あり学習の手法の一つであり、知識ベースから取得した少数の手がかり表現を用いることで、半自動的な教師データの収集を可能にする学習方法である。M. Mintz [7] らは Web テキストからの関係性抽出のタスクに Distant Supervision の考え方を適用し、少量の知識ベースから大量の学習データを収集している。この際、知識ベースとして FreeBase を用いている。三浦ら [8] は、Twitter の投稿に対するセンチメント分析のタスクに Distant Supervision の考え方を適用することで、教師あり機械学習の低コスト化に成功している。この際、学習データ収集に用いる手がかり表現として顔文字を用いている。山本ら [9] は、Web 上のニュース記事からの企業間関係抽出のタスクに Distant Supervision を適用している。ここでは、あらかじめ作成した教師データから、その判断の決め手となった語を抽出し、手がかり表現としている。これらの研究では Web 上からの情報抽出に関するタスクにおいて、分類性能を低下させることなく、教師あり機械学習の低コスト化に成功している。これらの結果から、適切な手がかり表現の定義が可能であれば、情報抽出のタスクにおいて Distant Supervision の考え方を適用することで、低コストで教師データを取得できると考えられる。

## 3. 本研究で扱う要望の定義

本研究の要望の定義は、大森と栗原らの論文を参考にした。大森は、直接的な要求を表す文章は、「命令」、「依頼」、「禁止」、「誘いかけ」、「希望」、「当為」、「当为非断定」、「希望非断定」のいずれかの態度を帯びるとし、それぞれの文法的な特徴を明らかにしている。また、栗原らは大塚ら [10] の研究を参考に、「～してほしい」「～てください」「～てくれ」といった、日本語母語話者のほとんどが「要求」と判断できる表現を「直接要求」表現とし、「～べき」「～がベストだと思う」「が必要」といった、「～してほしい」に言い換え可能な表現を「要求意図」表現とした。栗原らはさらに、Twitter 上の投稿は自由回答アンケートと異なりユーザの独り言や愚痴が投稿される傾向があることに注目し、直接要求や要求意図に当てはまらない場合でも、その内容が要望の動機になる否定的なテキストを「不満」と定義し、「直接要求」「要求意図」「不満」の3つに該当するテキストを要望と定義した。

本研究では以上の先行研究を踏まえ「命令」、「依頼」、「禁止」、「誘いかけ」、「希望」、「当為」、「当为非断定」、「希望非断定」の態度を帯びる表現と、これらに該当しないが要望の動機となる否定的な表現を「不満」とし、まとめて要望と定義する。以下、それぞれの表現について具体例を含め整理する。

### 命令

相手が意志的に制御できる動作を、相手に強制する表現

例 1) つまらん心配はしないで早く行け

例 2) はやくバグ修正しろ

### 依頼

相手の意志を尊重して、相手にある動作をするよう頼む表現

例 3) あなたはやく帰ってきてちょうだい

例 4) ちょっと、その婆さんに会ってみてくれないか?

### 禁止

相手にある動作をしないこと、あるいは、ある事態が生じないように努力することを命令する表現

例 5) そういうことに、やたら興味を持つな

例 6) いちいちアップデートすんな

## 誘いかけ

聞き手に、話し手と同様の行動をとるように要求する表現

例 7) やりましょう, 松田さん熊谷さん

例 8) 一緒にゲームしましょう!

## 希望

話し手自身に関わる事態の実現を希望する, あるいは他者がある事態を実現することを希望する表現

例 9) 千葉へ行ってもらいたい

例 10) 早く返金してほしい

## 希望非断定

希望の態度を断定することを控える表現

例 11) 音楽というコンテンツを手に入れたら, 通勤の時に電車で iPod やその他携帯音楽プレイヤーで聴きたいかもしれない

例 12) Windows も Mac も辞書データをひっくるめて月額制でお安くしておきますよという, プレミアムコースを作ってもらいたいかもしれない

## 当為

ある事態が望ましいとか, 必要だ, というように事態の当否を述べる当為の態度のうち, 「～べきだ」, 「～なければならぬ」のような述語の基本形をとって表される表現

例 13) 日本は早急に貿易黒字を減らすべきだ

例 14) 君は, あの時彼と別れるべきだった

## 当為非断定

当為の態度を断定することを控える表現

例 15) 日本は早急に貿易黒字を減らすべきだろう

例 16) 君は積極的になったほうがいいのかもしい

## 不満

「命令」, 「依頼」, 「禁止」, 「誘いかけ」, 「希望」, 「当為」, 「当為非断定」, 「希望非断定」に該当しないが, 要望の動機となる否定的な表現

例 17) 横浜市営地下鉄の始発遅い, 最悪

例 18) 市役所の対応悪いわ

## 4. 提案手法

### 4.1 提案手法の概要

本研究における提案手法の枠組みを図 1 に示す。まず, 商品名/サービス名を含むツイートの収集を行う。次に収集したツイートに対して後述する要望表現辞書と n-gram 判定の 2 段階の処理によって教師データの抽出を行う。その後, 教師データから素性の構築と学習を行い, 分類器を生成する。

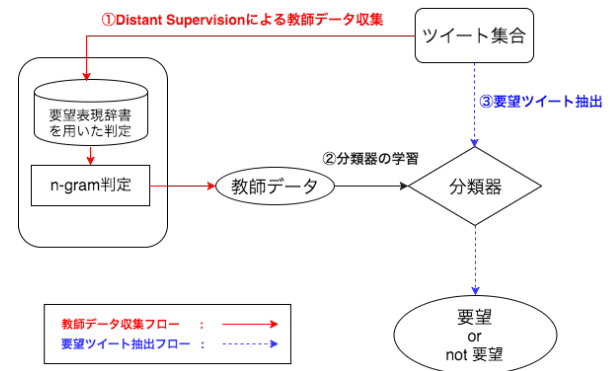


図 1 提案手法の概要

### 4.2 要望表現辞書の作成

Distant Supervision の手法を用いた教師データの収集では, まず始めに, 教師データの判別の手がかりとなる表現を決定する。それらの表現的な特徴を含むデータを収集することによって, 半自動的な教師データの収集が可能になる。したがって, Distant Supervision の手法を用いて教師データを収集するには, まず, 教師データの特徴を決定する必要がある。以降, 本論文では, 教師データの特徴となる表現のことを手がかり表現と呼称する。本研究では, 大森らの論文に記述されている要望文の特徴表現リストを参考に, 合計 19 個の特徴的な表現を定義し, 手がかり表現とする。これらの手がかり表現を要望表現辞書とした。

### 4.3 教師データの収集

4.2 節で作成した要望表現辞書を用いて, 教師データを収集する。その際に重要となるのは, 手がかり表現の出現位置である。要望表現辞書に定義した手がかり表現は文末に出現する傾向が極めて高く, 文末以外で出現する場合は, 要望とはならない可能性が高い。そこで本研究では, より高い精度で教師データを収集する為に, 手がかり表現の出現位置を考慮した, 2 段階の処理で教師データの収集を行う。

### 4.3.1 要望表現を含むツイートの抽出

4.2節で作成した要望表現辞書を用いて、TwitterAPIから対象の商品名/サービス名を含み、かつ手がかり表現を含むツイートを抽出する。この際、特定の語が出現するツイートをスパム・ボットによる投稿と判断して除去している。

### 4.3.2 要望表現が文末から 3-gram 以内に出現するツイートの抽出

先行研究から、要望表現の特徴は、文末に出現する頻度が高いことが知られている。そこで、本研究では、4.3.1節で抽出したツイートを文単位に分割し、文末から 3-gram 以内に手がかり表現が出現するものを選択し、教師データとした。ツイートの文単位への分割には、句点・空白・顔文字といった表現を区切り文字として使用し、これらの表現が出現した位置でツイートを分割し、1文と見なしている。また、3-gram 以内に出現するかの判定には、形態素解析器である Mecab を使用した。MeCab を用いてツイートを形態素に分解し、手がかり表現が、文末の形態素から 3 形態素以内の距離に存在しているかを確認することで判断している。

## 4.4 分類器の生成

4.3節で収集した教師データを用いて分類器の構築を行う。分類器のアルゴリズムには情報抽出のタスクにおいて有効性が知られている SVM を用い、実装には Python の機械学習ライブラリである scikit-learn を使用した。

素性には、単語の出現頻度などによって文書をベクトルで表現する形式である Bag of Words(BoW) を用いる。ただし、品詞が、名詞・動詞・形容詞・形容動詞・副詞・助動詞のいずれかに該当しない単語は除外した。また、素性の構築時には 文章内において一定以上の割合で出現する単語を頻出語として除去している。

## 5. 評価実験

### 5.1 評価対象

ソーシャルゲームは、近年急速な普及を見せており、多くの要望が Twitter 上に投稿されている。そこで本研究ではソーシャルゲームに関する投稿を評価対象とした。また、分類器を構築する際に使用する教師データとして、主要ソーシャルゲーム 10 タイトルに関する投稿を使用する。この際、要望を含むツイートを 2 万件、非要望ツイートを 1.5 万件、合計 3.5 万件のツイートを学習データとして収集している。

評価用データは、2015 年 8 月から 2 ヶ月間に投稿されたツイートを以下の 2 通りの方法で収集した。

### データセット (i) 教師データで使ったゲームタイトルを使用する

教師データとして使用したソーシャルゲーム 10 タイトルのタイトル名を含むツイートを各 200 件ずつ、合計 1000 件のツイートを収集した。この際、ゲームの公式 Twitter アカウントに対するリプライも収集対象としている。結果、要望を含む投稿は 221 件得られた。

### データセット (ii) 教師データで使ったゲームタイトルを使用しない

教師データとして使用していないソーシャルゲームタイトル「白猫プロジェクト (白プロ, #白猫)」を含むツイートを 1000 件収集した。この際、ゲームの公式 Twitter アカウント (@wcat\_project) に対するリプライも収集対象としている。結果、要望を含む投稿は 167 件得られた。

以上の手法により収集したツイートに対して、クラウドソーシングサービスのランサーズを用いて、要望のラベルを手手で付与した。各ツイート毎に 5 名の参加者に回答してもらい、もっとも一致率の高い解答を正解ラベルとして付与した。この際、回答の質を向上させる為に、100 ツイート毎に解答難度の低いダミーデータを用意し、ダミーデータへの回答を誤った参加者の解答を事前に除去している。回答者の判別の一致度を示す  $\kappa$  係数は、データセット (i) で 0.468, データセット (ii) で 0.548 となり、5 人の解答はおおむね一致していることが分かる。本研究では、以上の手順により作成したラベル付きツイートを評価用データとして使用する。

### 5.2 評価方法

提案手法と従来手法の分類精度の比較を行うことによって、本手法の有効性を検証する。提案手法の有効性を検証するには、抽出したツイート集合がどれだけ正解しているかという正確性と、抽出した記事集合が全ての正解のうち、どれだけ正解を含んでいるかという網羅性の 2 つの観点からの評価が必要となる。本論文では、正確性を適合率 (precision), 網羅性を再現率 (recall), 適合率と再現率の調和平均である F 値 (F-measure) によって提案手法の抽出精度を評価する。それぞれの計算方法について、以下に示す。

$$precision = \frac{\text{抽出した正解ツイート数}}{\text{抽出したツイート数}} \quad (1)$$

$$recall = \frac{\text{抽出した正解ツイート数}}{\text{全ての正解ツイート数}} \quad (2)$$

$$F\text{-measure} = \frac{2 \times precision \times recall}{precision + recall} \quad (3)$$

### 5.3 実験設定

#### 5.3.1 SVM の設定

SVM のカーネルには線形カーネルを使用した。また、パラメータの設定では、C パラメータの値を 1.0 に設定している。

#### 5.3.2 従来手法の実装

比較対象である辞書ベース分類器の実装に当たっては、栗原ら [4] の論文を参考に、大塚ら [10] の論文中に記述されている要望表現リストを辞書の作成に使用した。データセット (i)、データセット (ii) の手順により収集した評価用データに対して、作成した辞書内の要望表現とマッチするかどうかの判定を行い、一致する場合に要望を含む投稿であると判定する。

### 5.4 実験結果・考察

実験結果を、表 1 に示す。表 1 よりいずれの手法の評価データを用いた場合でも、従来手法と比較して、適合率、再現率、F 値が向上していることが分かる。特に再現率は、評価指標の中でも大幅な精度向上を確認することが出来た。この要因としては、Distant Supervision の手法を用いて収集した大量の学習データが、Twitter 上の多様な文章表現への対応を可能にした為であると考えられる。

また、データセット (i)、データセット (ii) を比較すると、教師データの収集時に使用していない未知のゲームタイトルに関して、教師データの収集時と同様のゲームタイトルを使用したデータセット (i) と同等以上の分類性能を発揮していることが確認できる。逆に、データセット (i) よりもデータセット (ii) の場合に高い分類性能を発揮している。この要因としては、データセット (i)、データセット (ii) で収集したツイートの判別難易度の違いが考えられる。評価データのラベル付けを行った際の  $\kappa$  係数の値は、データセット (i) で 0.468、データセット (ii) で 0.548 となっており、データセット (i) で評価者の回答一致率が低いことが確認できる。このことからデータセット (i) の評価データは、人手でも判別の難易度の高いツイートを多く含んでいる可能性がある。

また、適合率に関しては、従来手法と比較して、大幅な精度向上を達成することは出来なかった。この要因としては、まず、要望の対象が異なるツイートの判別難易度の高さが考えられる。今回の実験の抽出対象は、3章で挙げた条件に加えて、対象のソーシャルゲームに関する要望であるという 2つの要件を満たすツイートに限定している。しかしながら、従来のテキストマイニングで用いられる BoW 素性は、単語の出現頻度を特徴量とする為、単語の出現順序や係り受け関係を考慮することが困難である。これにより、分類精度が大きく低下してしまう可能性が高い。可次に考えられる要因としては、Twitter 上の要望の人手による

判別難易度の高さが挙げられる。今回使用した評価データセットは、いずれの手法を用いて収集した場合でも、5人の回答が一致しないケースが多く見受けられた。Twitter 上の投稿は、文章が省略されていたり、日本語が崩れていることが多い。加えて本研究で要望の対象としている不満に関するツイートは、暗にユーザの要望を示すあいまいな表現が多い為、より判別を困難にしていることが推測される。人手でも評価が一致しないツイートを抽出する事は困難である為、分類精度を低下させている可能性が高い。

表 1 実験結果

		適合率	再現率	F 値
データセット (i)	提案手法	0.24	0.42	0.31
	従来手法	0.2	0.06	0.12
データセット (ii)	提案手法	0.29	0.62	0.40
	従来手法	0.19	0.09	0.12

### 5.5 追加実験

実験結果より得られた仮説を検証する為に、追加で以下の実験を行った。

#### 5.5.1 「要望の対象が異なる」ツイートを要望に含めた場合との精度変化を比較

要望の対象が異なるツイートの判別難易度の高さが分類精度に与える影響を検証する為に、要望の対象が異なるツイートを抽出対象に含めた場合と含めなかった場合で、分類精度の変化を分析した。表 2 より、いずれの方法で収集した評価データに対しても、要望の対象が異なるツイートを抽出対象に含めた場合に、適合率・F 値が上昇していることが確認できる。この要因としては、分類器が、要望の対象が異なるツイートを、要望を含むツイートとして誤判定してしまっている事が考えられる。要望の対象を考慮した上で、より高い精度での分類を実現する為には、単語の出現順序や、係り受け関係などを考慮した、新たな素性を追加する必要がある。

表 2 要望の対象が異なるツイートを要望に含めた場合の分類結果

		適合率	再現率	F 値
データセット (i)	含める	0.36	0.42	0.39
	含めない	0.24	0.42	0.31
データセット (ii)	含める	0.51	0.62	0.56
	含めない	0.29	0.62	0.40

#### 5.5.2 回答一致率毎の精度変化を比較

人手による判別の難易度が分類精度に与える影響を検証する為に、回答の一致数毎に評価データを用いて分類を行った。実験結果を図 2、図 3 に示す。図 2、図 3 からいずれの手法で収集した評価データにおいても、回答の一致数が 5 の時に最も F 値が高くなり、一致数が 3 の時に最も F 値が低くなる事が分かる。この結果より、人手による

判別の難易度の高さが、分類精度に影響を与えていることが確認できる。また、一致数が5のツイートのみを使用した場合と一致数が3のツイートのみを使用した場合の分類精度の差を比較すると、データセット (i) では、-0.07、手法 (ii) では -0.14 となっており、極端な精度の低下は見られなかった。以上の結果より、本手法の人手による判別難易度の高いデータに対するロバスト性が確認できる。

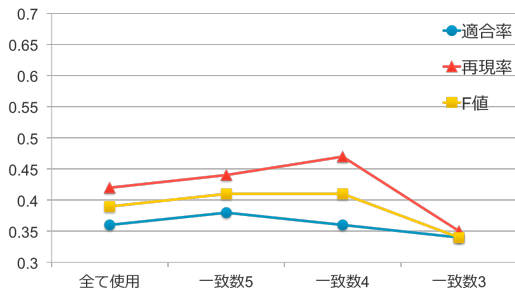


図 2 データセット (i) を対象とした回答一致率毎の精度変化

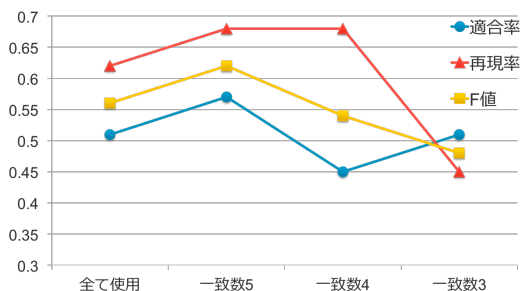


図 3 データセット (ii) を対象とした回答一致率毎の精度変化

## 6. おわりに

本論文では、Twitter 上からより高い精度で要望に関する投稿を抽出することを目的に、Twitter から消費者の要望を含む投稿の抽出手法を提案した。本手法では、要望表現の抽出に教師あり機械学習のアルゴリズムである SVM を用いることで、従来手法と比較して、より高い精度での要望抽出に取り組んだ。また、教師データの収集に半教師あり学習の一つである Distant Supervision を適用することで、低コストな機械学習の実現を試みた。

評価実験では、ソーシャルゲームに関する投稿を対象とし、提案手法と従来手法の分類精度の比較を行うことによって、本手法の有効性を検証した。教師データの収集時に使用したゲームタイトル名を使用する場合と、使用しない場合の2通りの方法でソーシャルゲームに関する投稿を各 1000 件ずつ収集し、人手でラベル付けを行ったものを評価用データとして用意した。評価データに対して、構築した分類器を用いて分類を行った結果、いずれの方法で収集した評価用データに対しても、適合率、再現率、F 値において提案手法が高い評価を示し、有効性を確認できた。

追加実験により、要望の対象が異なるツイートの判別難易度の高さが、分類精度にネガティブな影響を与えていることが検証された。また、回答一致率毎の精度変化を比較することで、人手による判別難易度の高い問題に対する本手法のロバスト性が確認できた。今後の課題としては、単語の出現順や序や係り受け関係を考慮した、新たな素性の追加が挙げられる。

## 謝辞

本研究は、NII 戦略研究公募型共同研究ならびに JSPS 科研費 25280110 の助成を受けたものです。ここに記して謝意を示します。

## 参考文献

- [1] 萩原雅之：“次世代マーケティングリサーチ”，ソフトバンククリエイティブ (2011).
- [2] 奥村学：“マイクロブログマイニングの現在”，電子情報通信学会技術研究報告. NLC, 言語理解とコミュニケーション, **111**, 427, pp. 19–24 (2012).
- [3] 野畑周, 内藤弘朗, 清水徹：“ヤフージャパンのリアルタイム検索における感情分析”，電子情報通信学会技術研究報告. NLC, 言語理解とコミュニケーション, **114**, 211, pp. 31–35 (2014).
- [4] 栗原理聡, 佐々木彬, 松田耕史, 岡崎直観, 乾健太郎：“Twitter を利用した地域ごとの要望抽出”，第 29 回人工知能学会全国大会, pp. 1–4 (2015).
- [5] 山本修平, 佐藤哲司：“Twitter からの実生活情報の抽出法の提案”，第 4 回データ工学と情報マネジメントに関するフォーラム (DEIM Forum 2012), F3-4 (2012).
- [6] 大森晃：“要求抽出のための言語学的基礎論：要求概念の定義, および要求の態度”，情報科学技術フォーラム講演論文集, **8**, 2, pp. 167–174 (2009).
- [7] M. Mintz, S. Bills, R. Snow and D. Jurafsky: “Distant supervision for relation extraction without labeled data”, Proceedings of the Joint Conference of the 47th Annual Meeting of the ACL and the 4th International Joint Conference on Natural Language Processing of the AFNLP: Volume 2 - Volume 2, ACL '09, Stroudsburg, PA, USA, Association for Computational Linguistics, pp. 1003–1011 (2009).
- [8] 三浦康秀, 服部圭悟, 大熊智子, 増市博：“Distant supervision による感性トピックの抽出”，富士ゼロックス テクニカルレポート, **23**, pp. 72–80 (2014).
- [9] 山本彩奈, 宮村祐一, 中田康太, 岡本昌之：“Deepdive を用いた web ニュース記事からの企業間関係抽出”，DICOMO シンポジウム 2015, pp. 172–179 (2015).
- [10] 大塚裕子, 内山将夫, 井佐原均：“自由回答アンケートにおける要求意図判定基準”，自然言語処理, **11**, 2, pp. 21–66 (2004).