

感情強度辞書を利用した日本語ツイートの感情強度の推定 Estimating Emotion Intensities in Japanese Tweets Using Emotion Intensity Lexicon

赤堀 立樹^{*1}
Tatsuki Akahori

堂坂 浩二^{*2}
Kohji Dohsaka

伊東 嗣功^{*2}
Hidekatsu Ito

石井 雅樹^{*2}
Masaki Ishii

^{*1} 秋田県立大学 大学院
Graduate School, Akita Prefectural University

^{*2} 秋田県立大学
Akita Prefectural University

This paper presents a method for estimating emotion intensities in Japanese tweets using emotion intensity lexicon. We intend to utilize this method to create dialogue systems that can regulate user's emotion. Such dialogue systems can be expected to be useful for maintaining the good mental condition and interpersonal relationships of the users. We first created a crowdsourced emotion intensity lexicon for eight categories of emotion (joy, sadness, anger, fear, trust, disgust, anticipation, surprise), which captures word-emotion intensities using best-worst scaling. To the best of our knowledge, there is no previous works to construct emotion intensity lexicon in Japanese. Then we carried out an experiment to evaluate the effect of the emotion intensity lexicon in the estimation of emotion intensities in Japanese tweets. The experimental results show that the emotion intensity lexicon can improve the accuracy of estimating emotion intensities in tweets.

1. はじめに

対話システムには、ユーザとの対話を通して道案内や予約など特定のタスクを遂行するタスク指向型対話システムと、雑談などコミュニケーション自身を目的とする非タスク指向型対話システム(雑談対話システムとも呼ぶ)があり[中野 2015]。近年は雑談対話システムの研究が盛んに行われている。人間同士の雑談において、対話参加者は相手の感情を考慮し、相手の感情を調整(促進・維持・抑制)しながら対話に従事する。例えば、悲しんでいる人に対して慰める発話をするにより、相手の「悲しみ」を和らげたり、「喜び」の感情をもっている人に対して賞賛や共感するような発話をするにより、相手の「喜び」を促進する。雑談対話システムにおいても、ユーザの感情を考慮し、ユーザの感情を調節することができれば、ユーザの良好な心身状態や対人関係の維持、また感情調節困難な方の支援に貢献できることが期待される。

ユーザの感情を調節する対話システムを構築するためには、ユーザの感情強度やその変化を測定することが必要である。感情強度の推定に関する研究は盛んになりつつあり、WASSA-2017 [Mohammad 2017] や SemEval 2018: Task 1 Affect in Tweets [Mohammad 2018a] では、4感情(anger, sadness, fear, joy)を対象として、Twitter における英語ツイートの感情強度を推定するタスクが設定されており、様々な手法が提案されている。これらの研究は英語の感情強度推定が目的であるが、その一方、佐藤らは日本語ツイートの感情強度を推定する手法を提案している[佐藤 2018]。

ツイートの感情強度推定のためには、感情辞書が重要な役割を果たす。英語の感情強度推定の場合、英語の言語表現に対して感情強度を付与した感情強度辞書 [Mohammad 2018b] が利用されている。日本語の感情辞書としては、日本語表現を10種類の感情に分類した感情表現辞典[中村 1993]や単語に感情極性(positive・negative)の値を実数値[-1.0, 1.0]で付与した単語感情極性対応表[高村 2006]などがあるが、管見の限りでは日本語の感情強度辞書は構築されてこなかった。

そこで、本研究では、日本語の感情強度辞書を構築し、感情

表 1. 感情分類

Plutchik の 8 感情分類	中村の 10 感情分類
喜び	喜
受容	好・安
嫌悪	厭
悲しみ	哀
怒り	怒
恐れ	怖
予期	NRC-Emotion-Lexicon
驚き	驚

強度辞書を用いた日本語ツイートの感情強度の推定法を評価することを目的とする。以下において、第 2 節で日本語感情辞書の構築について説明する。第 3 節で感情強度辞書を使った日本語ツイートの感情強度推定の評価実験について述べ、第 4 節で実験結果について考察する。

2. 日本語感情強度辞書の構築

本研究では、まず、感情強度辞書に登録する感情表現を収集する。その感情表現の強度を数値化し、感情強度辞書を構築する。

2.1 感情表現の収集

本研究では、中村の「感情表現辞典」[中村 1993] から感情カテゴリ(喜, 好, 安, 厭, 哀, 怒, 怖, 驚, 恥, 昂)ごとに感情表現を収集し、中村の 10 分類を Plutchik [Plutchik 2001] の 8 分類(喜び, 受容, 嫌悪, 悲しみ, 怒り, 恐れ, 予期, 驚き)に再分類した。再分類は他研究との比較や応用しやすくするためである。表 1 に再分類の振り分けを示す。ここで、中村の感情分類の「好」と「昂」は、他の複数の感情に分類できるため、適宜分類した。例えば、「昂」の「飛び上がる」や「踊る」などは「喜び」に分類し、「熱り立つ」や「怒り狂う」は「怒り」に分類した。また、Plutchik の感情分類の「予期」は中村の感情分類には該当する感情がないため、Plutchik の感情分類に対応する言語表現を人手により収集した。

表 2. 感情表現 BWS 結果(喜び, 受容, 嫌悪, 悲しみ)

喜び		受容		嫌悪		悲しみ	
嬉し泣き	0.979	信頼する	0.944	反吐が出る	0.924	悲痛な叫び	1.000
歓喜	0.974	敬愛する	0.875	目障り	0.905	咽び泣き	0.984
天にも上る気持ち	0.974	相思相愛	0.854	生理的に無理	0.896	泣き崩れる	0.923
大喜び	0.970	信仰する	0.854	殺意を抱く	0.894	泣き叫ぶ	0.875
狂喜	0.917	尊ぶ	0.852	吐き気がする	0.875	見るのもつらい	0.875
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
おかしい	0.119	ゆるゆる	0.143	羨ましい	0.167	しける	0.109
笑	0.119	のどか	0.119	困惑	0.146	くすんくすん	0.078
気軽	0.095	無事	0.111	残念に思う	0.125	デメリット	0.047
吉	0.083	ふかふか	0.104	無念	0.095	苦手	0.042
クスクス	0.083	ゆったり	0.097	可哀想	0.000	ガーン	0.000

表 3. 感情表現 BWS 結果(怒り, 恐れ, 予期, 驚き)

怒り		恐れ		予期		驚き	
ぶっ殺す	1.000	戦慄	0.938	予兆	0.938	言葉が出ない	0.929
殺す	0.984	叫喚	0.938	予測する	0.917	言葉を失った	0.881
殺意	0.911	阿鼻叫喚	0.913	兆候	0.917	想像を絶する	0.871
怒鳴り散らす	0.909	生きた心地がしなかった	0.859	兆し	0.905	夢にも思わない	0.854
怒り狂う	0.900	畏怖	0.859	予想する	0.833	驚愕する	0.833
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
頬を膨らます	0.141	渋る	0.125	宿命	0.208	ビクビク	0.185
くしゃくしゃ	0.141	決めかねる	0.113	釈然としない	0.208	認められない	0.167
つまらん	0.125	期待できない	0.094	理想	0.200	たまたま	0.042
ふくれっ面	0.089	気遣い	0.094	大志	0.167	うつとり	0.042
つまらない	0.075	堂々巡り	0.094	ドリーム	0.167	はわわ	0.024

NRC-Emotion-Lexicon [Mohammad 2013] を日本語に翻訳した表現集を使った。

次に、多様な表現を収集するために、収集した感情表現の類語を収集した。すべての感情表現の類語を収集するのは、コストがかかるため、Twitter API を使って 2017 年 3 月から 2018 年 9 月までの間に収集した約 2000 万ツイートの上で出現回数を調べ、感情カテゴリごとに出現回数上位 30 語程度の感情表現の類語を収集することとした。類語の収集には Weblio 類語辞典¹を使用した。類語の収集の際には、明らかに感情表現と意味合いが異なる類語は除外した(例:「怒り」の感情表現「雷」から稲妻・電光などの類語は収集しない)。類語を持たない語(例:爆発する)も除いた。類似した類語(例:激怒、憤怒、憤慨)は 1 つを残して除いた。最後に、類語を含めた感情表現の出現回数を再度調べ、出現回数 1 回以上の語を抽出した。

2.2 節で説明するように、本研究では、Best-Worst Scaling (BWS) [Louviere 1991] を用いて、感情表現に感情強度を付与する。BWS は相対評価法であるため、感情強度に極端な偏りがあると適切に算出できない。そのため、抽出した感情表現の強さを手動で 3 段階に分類し、収集した感情表現の強度に極端な偏りが無いことを確認した。感情強度の 3 段階の分類は、筆者の一人が各表現に対して 0~100 の実数値を付与することによって行った。3 段階は大(68~100)、中(34~67)、小(0~33)とした。以上の結果、1051 語の感情表現を収集した。

2.2 感情表現の強度の数値化

2.1 で収集した各感情表現に対して BWS によって感情強度を数値化した。BWS は Mohammad らの手法 [Mohammad

2018b] を参考に行った。感情カテゴリごとに 4 人の作業者に $2 \times N$ (N は感情表現数)の質問に回答してもらった。各質問では、特定の感情カテゴリに属する 4 つの感情表現が与え、最も感情を強く感じる感情表現 (Best) と最も感情を弱く感じる感情表現 (Worst) の 2 つを回答することを求めた。このとき、Best と Worst は異なる感情表現を選択してもらった。各感情表現は $2 \times N$ の質問中に異なる質問で 8 回出現し、同一の選択肢の質問は現れないようにした。感情表現 e の感情強度の値は次の式によって計算する。

$$\frac{P_{Best}(e) - P_{Worst}(e) + 1}{2}$$

このとき、 $P_{Best}(e)$ 、 $P_{Worst}(e)$ は感情表現 e がそれぞれ Best、Worst に選択された確率である。 $P_{Best}(e)$ 、 $P_{Worst}(e)$ の値域は $[0, 1]$ であり、 $P_{Best}(e) - P_{Worst}(e)$ の値域は $[-1, 1]$ となる。本研究では値域を $[0, 1]$ とするために線形変換を行う。感情強度は値が 1 に近いほど強い感情を示し、0 に近いほど弱い感情を示す。

BWS はクラウドソーシング Lancers を使って実施した。作業の質を確保するために全質問の 5% の質問を回答が明らかに一意に定まる質問 (gold question) として設定し、gold question の正解率が 70% 未満の作業者の回答は使用しなかった。gold question は、BWS の作業者とは別の 2 名の作業者に感情強度「大」「小」からそれぞれ 1 つと「中」から 2 つから成る質問に対して BWS と同様に回答してもらい、2 人の回答が一致した回答の中から作成した。

BWS は単一の質問で複数の項目間の順序関係を得られるため、簡便かつ有効な評価方法であると考えられる。表 2, 3 に各感情の感情強度の値が上位と下位それぞれ 5 つずつの結果を示す。

¹ <https://thesaurus.weblio.jp>

2.3 感情強度辞書に関する考察

2.2 節の方法により、1051 語からなる感情強度辞書の作成を行った。感情強度の値は概ね直感と一致する結果となったが、予期の感情では一部直感とは異なる結果が見られた。

予期の感情では、感情強度が高いと考えられる「期待」、「宿命」、「理想」といったの表現が強度 0.67 以下の中・下位に入った。一方、感情強度が低いと考えられる「知らせ」、「見込み」などの表現が、感情強度 0.67 以上の高位に入った。このようになった要因として、作業者が「予期」の感情の意味を限定的に捉えた可能性がある。予期の感情には「予想」、「知らせ」、「見込み」といった未来の予見や未知の情報を表す表現の他に、「宿命」といった決まった未来を表す表現や、「期待」、「理想」といった希望・意図を表す表現がある。作業者は予期の感情について、未来の予想や未知の情報を表す意味に限定して考えたため、その他の意味を表す表現の強度が相対的に低くなった可能性がある。

3. ツイートの感情強度の推定と評価

3.1 感情強度辞書を使った感情強度推定方法

本研究では、感情強度辞書を使ったツイートの感情強度推定法の開発の第一歩として、簡易な方法により感情強度の推定を行い、感情表現辞書の効果について考察する。

前提として、ツイートの感情分類が成されているものとした。また、感情強度辞書に登録された感情表現を含むツイートを対象とした。ツイート中に含まれる否定や仮定、強調表現 (!! , オオオオオなど) は考慮しなかった。

ツイートの感情強度の推定は、感情強度辞書を用いて、ツイートに含まれる感情表現を手がかりとして行う。ツイートには複数の感情表現が含まれる場合もあるため、以下に示す、First, Last, Average の 3 つの方法と、ランダムに感情強度値を生成する方法の計 4 つの手法により推定を行う。各推定方法について以下に示す。

- First : ツイート中で最初に現れた感情表現の感情強度値をツイートの感情強度値とする
- Last : ツイート中で最後に現れた感情表現の感情強度値をツイートの感情強度値とする
- Average: ツイート中の感情表現の感情強度値の平均をツイートの感情強度値とする
- Random: ランダムに感情強度値を生成し、ツイートの感情強度値とする

3.2 評価実験

感情強度辞書を使ったツイートの感情強度推定の評価実験を行った。評価の対象となるツイートは「喜び」と「悲しみ」の感情に事前に分類された各 50 ツイートとした。ツイートの感情強度値の正解データを作成する作業者の負担を考慮して、ツイートは 20~50 文字程度のツイートのみを使用した。

ツイートの感情強度値の正解データ (測定値) を作成するため、事前に 3.1 節と同様の手順によりツイートの感情強度値を BWS により測定した。gold question の正解率が 70% の作業者の BWS の結果を使用し、正解データを作成した。

評価は、各方法により推定された感情強度値を大(0.67~1), 中(0.33~0.67), 小(0~0.33)の 3 段階への分類した際の正解率を使って行った。

表 4, 5 にツイートの感情強度推定の評価結果を示す。「喜び」の感情では、最初の感情表現の感情強度をツイートの感情

表 4. ツイートの感情強度推定評価結果(「喜び」)

	小	中	大	全体
a) First	60%	59%	69%	62%
b) Last	56%	50%	56%	54%
c) Average	60%	55%	60%	58%
d) Random	27%	48%	26%	34%

表 5. ツイートの感情強度推定評価結果(「悲しみ」)

	小	中	大	全体
a) First	37%	56%	54%	48%
b) Last	44%	56%	50%	50%
c) Average	44%	57%	54%	52%
d) Random	30%	47%	29%	35%

強度にする手法が正解率 62%で最も高い結果となった。また「悲しみ」の感情では、ツイート中の感情表現の感情強度の平均をツイートの感情強度とする手法が正解率 52%で最も高い結果となった。感情で評価結果を比較すると、「喜び」の感情のほうが全体的に高い正解率となった。

4. 考察

「喜び」の感情強度推定は 60%程度、「悲しみ」の感情強度推定は 50%程度の正解率となった。Random との比較ではあるが、感情強度辞書は一定の効果があると言える。「喜び」の感情の方が全体的に高い正解率となったが、作業者の感想から、「喜び」の感情に比べて「悲しみ」の感情は、感情表現の感情強度よりも、ツイート内容や顔文字の種類に影響されやすいことが考えられる。そのためツイートに含まれる感情表現のみを手がかりとする推定方法では「喜び」の感情の正解率が高くなったことが考えられる

表 6, 7 にそれぞれ「喜び」と「悲しみ」の感情強度推定が不正解となった場合の特徴的な例を示す。測定値とは、3.2 節で述べたように、事前に作業者の BWS により測定した感情強度の正解値である。推定値とは、感情強度辞書を使ってツイートの感情強度を推定した値である。「喜び」の場合は方法 First, 「悲しみ」の場合は方法 Average による推定値である。

まず、表 6 の「喜び」の場合、特徴的な例として、文中に否定を含む場合、辞書に含まれない感情を表す表現を含む場合、強調表現を多用している場合があった。まず、文中に否定を含む例として、表 6 のツイート番号 13 では、釣りに行ったが釣れなかったという否定内容が含まれていた。否定的な表現の影響

表 6. 感情強度推定の不正解例 (喜び)

ツイ ート 番号	ツイート内容	測定 値	推定 値
13	初めてうなぎつりやった てか、 昨晚は鯉つり w ミナミの釣り堀 〜 釣れなかったけど w 久々に 糸垂らして恍惚としてたよ	0.271	0.771
18	デザフェス当選した！！ヤッター ー！！！！！！	0.875	0.667
26	念願の！！！！チュロス(｡◕◕｡)この ベリーソース神ってる!! 店員 のお姉さんと、少しお話もして、 充実した朝時間になりました㇏ (●´▽`●)ﾉ	0.833	0.563

により、感情強度の測定値の方が、否定を考慮せずに推定された推定値よりも低くなったと考えられる。

ツイート番号 26 は、感情強度辞書に含まれていない感情表現を含む例である。「念願」や「神ってる」は辞書に含まれていないため、測定値の方が推定値よりも高くなったと考えられる。

ツイート番号 18 は強調表現「!!」を多用している例である。また、ツイート自体が短いため、強調表現による影響を受けやすい可能性がある。そのため、測定値の方が推定値よりも高い結果となったと考えられる。

表 7 に示す「悲しみ」の場合、不正解となった要因としては、顔文字の使用、ツイート内容、辞書に含まれない感情表現による影響が考えられる。まず、顔文字とツイート内容について考察する。表 7 のツイート番号 1 では感情強度の強い感情表現「咽び泣く」を使用しているが、ツイート内容または顔文字の影響により、悲しみの感情が弱く感じられたため、測定値が推定値より小さくなったと考えられる。一方、ツイート番号 12, 28, 32 では、測定値が推定値よりも高くなっており、顔文字が感情強度を高めている可能性がある。また、ツイート番号 36, 42, 45 でも、測定値が推定値よりも 0.2 程度高くなっているが、これらのツイートは、ツイートの内容の影響を受け、含まれる感情表現の感情強度値より高い測定値になった可能性がある。

次に、感情強度辞書に含まれない感情表現による影響について考察する。その例として表 17 のツイート番号 15, 47 がある。特にツイート番号 15 では、「阪神淡路大震災」がツイートに含まれている。現在作成した感情強度辞書は主に形容詞と動詞によって構成されている。今後は、感情強度値を付与した名詞を増やすことで推定性能が向上する可能性がある。

5. おわりに

本研究では、中村の「感情表現辞典」を基に類語の収集を行い、BWS により感情強度を測定し、1051 語からなる感情強度辞書を作成した。概ね直感と一致する辞書の作成を行うことができたが、予期の感情では一部直観と異なる結果となった。例として、予期の感情では、未来の予見や未知の情報を表す表現が上位の感情強度となり、決まった未来や希望・意図を表す表現が中・下位の感情強度となった。要因としては、作業者が予期の感情を未来の予想や未知の情報に限定して捉えたため、その他の意味をもつ表現の強度が相対的に低くなったと考えられる。感情強度辞書を使ったツイートの感情強度推定では喜び、悲しみともに 50～60% の正解率となった。感情強度辞書は一定の効果があることが示唆された。

今後の課題としては、感情強度値付きの大量のツイートからの学習による感情強度推定の性能向上、感情表現を含まないツイートの感情強度の推定が考えられる。また、名詞への感情強度値の付与や、否定や強調表現などを考慮した感情強度推定を行うことを検討する。

謝辞

本研究は科研費(16K00355)の助成を受けたものである。

参考文献

- [Louviere 1991] Jordan J. Louviere: Best-worst scaling: A model for the largest difference judgments, Working Paper, 1991.
 [Mohammad 2013] Saif M. Mohammad and Peter Turney: Crowdsourcing a Word-Emotion Association Lexicon, Computational Intelligence, Vol.29, No.3, pp.436-465, 2013.
 [Mohammad 2017] Saif M. Mohammad and Felipe Bravo-Marquez: WASSA-2017 shared task on emotion intensity. In:

表 7. 感情強度推定の不正解例(悲しみ)

ツイート番号	ツイート内容	測定値	推定値
1	なにによりあのぼっちりお目目と色白の肌とすっぴんとの髪が手に入るかと思うと色黒一重天然パーマな私は咽び泣きます(;▽;)ノ	0.396	0.984
12	まじっすか……(´ω`lll)寂しいです……やっぱ心痛みます…(´;ω`)	0.708	0.589
15	阪神淡路大震災から 23 年。被害に遭われた方々へ 哀悼の意を表します。この日を決して忘れません	0.771	0.672
28	災害よりもモリ・カケ、ですか。情けない。(ノ 口)。(ノ 口)	0.646	0.516
32	ミク、コヒメが卒業発表するシーンでコヒメから貰い泣きしてしまった(;;)無理いい(;;)	0.625	0.406
36	家で珈琲を入れるときにフレンチプレスってのていれてるんですけど、うっかりフィルターの紙を入れるのを忘れてた時があって、朝からがつくりしたことあります…	0.500	0.328
42	昔絡んでた方々は見えないところに行ってしまったよ(しんみり)	0.542	0.313
45	職場の人で「自分がやられて嫌なことを、自分は他の人にする派」がいてしょげる	0.500	0.268
47	自分の無知さを恥じる	0.583	0.234

Proc. the Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis, 2017.

[Mohammad 2018a] Saif M. Mohammad, Felipe Bravo-Marquez, Mohammad Salameh and Svetlana Kiritchenko: SemEval-2018 Task 1: Affect in Tweets, In: Proc. the 12th International Workshop on Semantic Evaluation, pp.1-17, 2018.

[Mohammad 2018b] Saif M. Mohammad: Word Affect Intensities, In Proc. LREC-2018, 2018.

[中村 1993] 中村明: 感情表現辞典, 東京堂出版, 1993.

[中野 2015] 中野幹生, 駒谷和範, 船越孝太郎, 中野有紀子, 奥村学 (監修): 対話システム, コロナ社, 2015.

[Plutchik 2001] Robert Plutchik: The nature of emotions, American Scientist, Vol.89, No.4, pp.344-350, 2001.

[佐藤 2018] 佐藤一輝, 尾崎知伸: 複数の表現学習手法を用いた日本語ツイートの感情強度推定, 人工知能学会研究会資料, SIG-KBS-B802-02, 2018.

[高村 2006] 高村大也, 乾孝司, 奥村学: スピンモデルによる単語の感情極性抽出, 情報処理学会論文誌ジャーナル, Vol.47, No.2, pp. 627-637, 2006.