

言語情報に基づく話し言葉の感情抽出に関する研究

Extraction emotion from spoken language based on the language information

薛 佳恵
Yoshie Setsu

康 鑫
Xin Kang

西出 俊
Shun Nishide

任 福継
Fuji Ren

徳島大学 工学部
Tokushima University, Faculty of Engineering

Spoken language contains various emotion information. While there are many ongoing researches on emotion recognition using linguistic information, few of them consider facemarks in their model. In this study, we incorporate the facemarks in an emotional expression dictionary to recognition emotion in Twitter messages.

1. はじめに

近年、人工知能技術の飛躍的進歩に伴い、ヒューマンマシンインタラクションが益々重要となった。ヒューマンマシンインタラクションでは、コンピュータによる感情識別が一つの目標である。

このような背景を受け、感情情報を直接に反映できる SNS も出現した。その体表例として、Twitter では、今起こったことや感じたことを気軽に投稿できる。これらの投稿では堅苦しい書き言葉よりも気楽な話し言葉を使って投稿することが多くみられる。そのため、Twitter の投稿にはユーザの感情が表れやすいと考えられる。従って、Twitter を使って感情情報を推定する研究が多く行われている。

SNS では、実際のコミュニケーションと異なり、ジェスチャーや表情、声の大きさ、トーン、スピードなどの非言語的な要素が含まれない。そのため、喜びの感情を表す(〇)のような顔文字を表情の代わりに使うことが多い。しかしながら、顔文字を生かした感情認識の研究はそれほどされていない。

三品ら[三品 10]は発話文を発話者の感情別に分類して構築した感情コーパスを用いることで感情推定を行う用例ベースの手法を提案した。実験を通して発話者の感情を推定することに成功している。この研究では一人の人間に対しての感情コーパスを構築し、さらに、その人の感情推定を行う。

山本ら[山本 14]は顔文字の役割を決定するために感情軸を感情極性に分類し、ユーザ実験を行った。その結果、文の感情極性、顔文字の感情極性、ツイートの感情極性から顔文字の役割を強調、弛緩、転換、付加の4つに分類した。

堀宮ら[堀宮 12]は Twitter ユーザの特性に関して、Twitter 上でフォローすべきユーザの推薦手法に関して、コミュニケーションに着目し研究した。ツイートの持つ感情ごとにそれに対するリプライにおける言語表現にパターンが存在するという仮説を立て、リプライ文を単語ベクトルとみなして、ツイートの感情推定を行う手法を提案した。

Renら[Ren 13] ツイートにおけるユーザトピック意見予測することに着目し、ScTcMF フレームワークを提案した。結果として、最先端の協調フィルタリング手法を実現し、ソーシャル文脈とトピック文脈が予測精度の改善に効果的であることを示した。

本研究では顔文字情報を生かした感情推定認識手法を提案する。つまり、言語情報と顔文字情報を統合した感情推定システムの構築である。なお、本論文では顔文字のみによる感情推定システムと言語情報のみによる感情推定システムを構築し、実験を通して、その二つの方法を比較し、有効性を検証した。

本研究では「喜び」、「愛」、「驚き」、「不安」、「怒り」、「悲しみ」の6種類の感情を考慮した。

2. ツイートの感情推定手法

図1のように、一つのツイート文に対して、顔文字と言語情報に分け、それぞれの感情辞書を用いて、感情語や顔文字に感情種類とその種類の感情値を付与し、合計値の一番大きい値の感情種類をその文の感情推定結果とする。

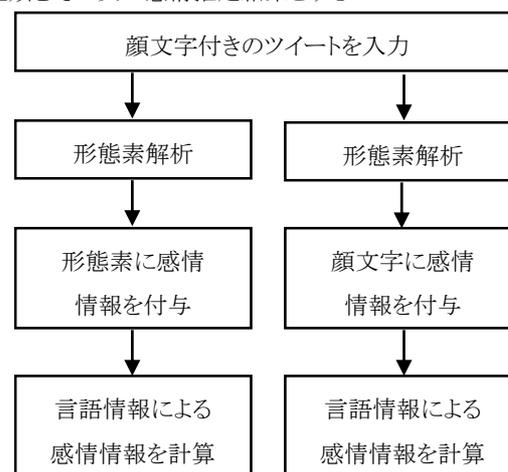


図1 感情推定手法

2.1 テキスト感情語辞書

テキスト感情辞書に各感情語に ID 番号と感情値が含まれている。作り方を以下に示す。

- 感情種類毎にテキストファイルに分割する
- 各テキストファイル内の文章を形態素解析にし、形態素の並びに変換する
- 各形態素について、感情語であるかを判断する
- 感情語である場合、出現頻度によって、重みを決定する
- 感情語ごとに ID 番号を割り振り、感情語の索引と感情語ごとの重みを出力する

2.2 顔文字の感情語辞書

本研究は 1000 個の顔文字を収集した。それぞれの、顔文字の感情値を決めるために、アンケートを実施した。

1000 個の顔文字をすべて一人に回答させることは負担が大きく、疲れによって不正確なデータとなる可能性がある。1000 個の顔文字を 100 個ずつの、10 グループにわけ、その中からラン

ダムに協力者に提示する。今回は 15 人から有効なデータを回収した。

アンケートでは、6 種類の感情に対して、一つの顔文字から感じた感情の程度を「強」、「中」、「弱」、「無」の4種類から選択してもらった。「強」に 0.9、「中」に 0.5、「弱」に 0.1、「無」に 0 と数値化し、それぞれの平均値をその顔文字の感情値とする。

3. 評価実験

Twitter から顔文字付きのツイートに 100 件取得し、被験者に各ツイートから感じた感情を「喜び」、「愛」、「驚き」、「不安」、「怒り」、「悲しみ」の 6 種類から一つを選択してもらい。選択された感情をそのツイートの正解感情とする。

その正解感情と顔文字のみによる感情推定の結果、言語情報のみによる感情推定の結果をそれぞれ表 1 と表 2 に示す。両システムの感情種類ごとに F 値を出し、比較した結果を図 2 に示す。

表 1 顔文字のみによる感情推定の結果

結果 正解	喜び	愛	驚き	不安	怒り	悲しみ	合計値
喜び	14	6	2	0	2	9	33
愛	12	5	2	0	0	3	22
驚き	10	0	2	1	0	1	14
不安	1	0	1	1	0	0	3
怒り	11	0	0	1	0	0	12
悲しみ	3	1	2	1	0	9	16
合計値	51	12	9	4	2	22	100

表 2 言語情報のみによる感情推定

結果 正解	喜び	愛	驚き	不安	怒り	悲しみ	合計値
喜び	15	6	4	2	2	4	33
愛	9	6	1	3	0	3	22
驚き	3	3	1	2	2	3	14
不安	1	0	1	0	1	0	3
怒り	2	1	3	4	2	0	12
悲しみ	2	1	0	2	0	11	16
合計値	32	17	10	13	7	21	100

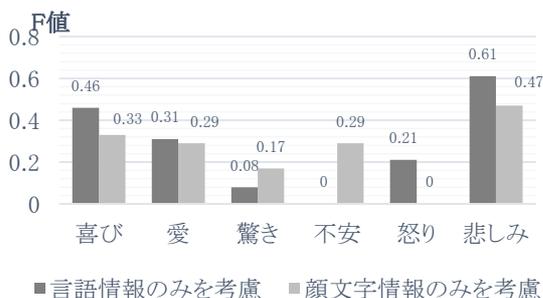


図 2 両システムの比較

図 2 より、顔文字のみを考慮した感情推定システムと言語情報のみを考慮した感情推定システムを比較した結果、全体的に顔文字のみの方が精度が劣っていることがわかる。感情種類ごとに詳しく見ると、「驚き」、「不安」の 2 種類の感情においては、顔文字のみを考慮した感情推定システムの方の精度が優れている。それ以外の感情種類においては、顔文字情報よりも言語情報のみを考慮した感情推定システムのほうの正解率が高い。

4. 考察

全体の精度において顔文字のみを考慮した感情推定システムの精度が劣る理由として、顔文字辞書の精度が低いことを考えられる。

「驚き」、「不安」の 2 種類の感情種類において、顔文字のみを考慮した感情推定システムの優れている理由として、この 2 種類の感情を表す感情表現が少ないことであると考えられる。一方で、顔文字の方が言語情報より直観的にわかりやすい。例えば、収集した顔文字付きのツイートの中の一文:「ほげー！？Σ(・艸・o)!!初課金武器ドロップ。ほんとに出るんだ...(´Д`)」がある。この文においては、感情表現のみをみると、驚きの表現があまり見当たらない。しかし、顔文字のみをみると、Σ(・艸・o)エと(´Д`)の二つとも驚きの感情を表す。

そのほかの感情種類では、言語情報のみを考慮した感情推定システムの正解率が高い理由として考えられるのは「驚き」、「不安」と違って、感情表現が多いことである。例えば、集めた顔文字付きのツイートの中で「ありがとう(つДつ)」や「カッコイイー!!Σ(・Д・)」という表現があったが、言語情報のみを考慮した場合は「喜び」になるが、顔文字のみを考慮した場合は「驚き」、「怒り」になる。このように、顔文字には曖昧さが存在しているために、言語情報のみで判断したほうが有効である場合もある。

5. おわりに

本研究では、顔文字のみによる感情推定と言語情報のみによる感情推定を比較した。結果として、全体的には顔文字のみによる感情推定の精度が劣っているが、「驚き」、「不安」の 2 種類の感情においては、顔文字のみを考慮した感情推定システムの方の精度が優れている。

今後の課題として以下の点が挙げられる。まずは、顔文字辞書の精度の向上である。より高精度な顔文字辞書を構築した後に、再度二つの方法を比較する。さらに、感情種類によって、顔文字のみと言語情報のみによる感情推定システムがそれぞれの有効性を示したため、両システムを統合した感情推定システムがより効果的になることが推定できる。二つの方法を統合した感情推定システムを構築することを目指す。

参考文献

- [山本 14] 山本湧輝, 熊本忠彦, 灘本明代: Twitter 特有表現を考慮したツイートの多次元感情抽出手法の提案, 情報処理学会関西支部支部大会, 2014.
- [三品 10] 三品賢一, 土屋誠司, 鈴木基之, 任福継: コーパスごとの類似度を考慮した用例に基づく感情推定手法の改善, 自然言語処理, Vol. 17, No. 4, pp. 91-110, 2010.
- [堀宮 12] 堀宮ありさ, 坂野遼平, 佐藤晴彦, 小山聡, 栗原正仁, 沼澤政信: Twitter における発話者へのリプライを用いたユーザ感情推定手法, DEIM Forum 2012, F2-1, 2012
- [Ren 13] Fuji Ren, Ye Wu, Predicting User-topic Opinions in Twitter with Social and Topical Context, IEEE Transactions on Affective Computing, Vol.4, No.4, pp.412-424, 2013