

リツイートしたユーザの表情に着目したリツイートのセンチメント分析

徳田 拓也[†] 横井 健[†]
東京都立産業技術高等専門学校[†]

1. はじめに

マイクロブログサービスの1つである Twitter[1]では、他のユーザの投稿(ツイート)を閲覧した際、何らかの思いや動機からそのツイートを他人と共有することができる。これはリツイートと呼ばれ、他のユーザの投稿を再投稿する機能である。しかしリツイートは他人の投稿の再投稿であるため、その表層表現からリツイートしたユーザのセンチメント情報を得られるとは限らない。

そこで本研究では、リツイート時の感情変化時に起こる表情の変化に着目し、リツイートしたユーザのセンチメント分析を行う。

2. 提案手法

本研究では、Web カメラを用いてリツイート時の表情を取得する。取得した表情に対して、顔認識 API[2]により画像処理を行い、感情の種類を検出する。この顔認識 API で判別された感情を表1の分類表を元にポジティブとネガティブの2種類に分類を行う。このポジティブとネガティブの分類を行った結果をリツイートされたツイートにセンチメント情報としてタグ付けをする。

また、笑顔強度の測定の実験で収集したツイートに対して形態素解析を行い、感情極性辞書を用いて感情極性を判別する。そしてこの判別結果と本手法の結果を比較し評価を行う。

形態素解析エンジンの MeCab[6] と MeCab 用の辞書に Unidic 辞書を使用し、ツイートに対して形態

素解析を行う。次に感情極性辞書を使用して解析された各単語の感情極性を調べ、ポジティブな単語は加算、ネガティブな単語は減算を行って合算された値を評価値として使用する。

感情極性辞書は単語感情極性対応表 [3]、日本語評価極性辞書(用言編) [4]、日本語評価極性辞書(名詞編) [5] の3種類を使用して評価を行う。

本研究では、リツイート閲覧時の表情のうち、特に笑顔の表情が出るツイートを対象とする。

3. 実験方法

5人の被験者に対してツイート閲覧時の笑顔強度の測定実験を行う。

Twitter 上の面白いと思われるツイートを31件収集し、被験者に評価してもらう。被験者は収集したツイートを作成したソフトウェア(図1)で順番に閲覧する。図1のツイート内容の部分に表示されたツイートが面白く、リツイートで他のユーザと共有したいと思った時は図1の「おもしろい」ボタンを押下する。同時に、Web カメラで被験者の表情をツイートの閲覧開始から閲覧終了まで、0.1秒間隔で取得する。全ツイートの閲覧終了後に、顔認識 API により画像処理を行い、笑顔の強度を測定する。

また、収集したツイートに対して形態素解析を行って感情極性を判別し、ツイート内容から推測される感情極性と表情から推測される感情極性を比較して評価を行う。

4. 実験結果

笑顔強度の測定の実験より取得した笑顔の強度データの散布図と、次数2の多項式近似による近似曲線を

表1 判別される感情の種類分類

ポジティブ	ネガティブ
喜び	怒り
落ち着いた	驚き
	困惑
	嫌悪

Sentiment Analysis of a Retweet with Facial Expression of a Retweeting User

[†] Takuya Tokuda

[†] Takeru Yokoi

[†] Tokyo Metropolitan College of Technology

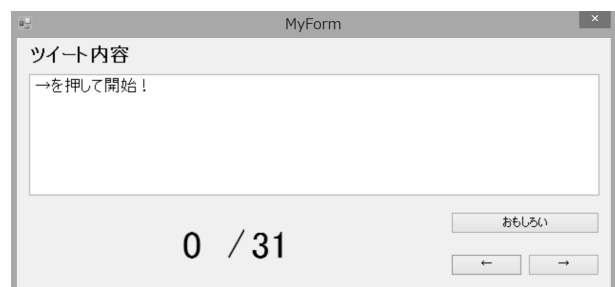


図1 作成したソフトウェアの外観

作成したグラフを図2に示す。図2は面白かった時にボタンを押下し、笑顔が強くなった被験者の笑顔の強度のグラフと、同被験者がボタンを押下しなかった時の笑顔の強度のグラフである。グラフ縦軸の笑顔の強度は0から100の百分率で表しており、横軸は経過時間を表している。また、図2の6.0[s]付近の縦破線はボタンが押下された時間を表している。

また、形態素解析を用いての感情極性判別の結果のうち、特徴があるものを表2に示す。IDは収集したツイートの通し番号である。単語感情極性対応表では、算出された評価値をMeCabで分割された単語の数で割った値である。日本語評価極性辞書では、ポジティブで一致した場合1を加算、ネガティブで一致した場合1を減算して評価値を算出している。面白いと思った被験者の数では、ツイート閲覧時に被験者が面白いと思いボタンを押下した人数である。

5. 考察

被験者がボタンを押下しなかった時に比べ、図2の面白かった時の笑顔の強度はボタン押下時に笑顔が強くなっていることがわかる。実験結果から、感情の変化が起こり表情に出てからボタンを押下しているの、常に表情を取得しなくともボタン押下時の表情を取得すれば表情からセンチメント情報を取り出すことができると考える。また、個人の顔つきや感情の変化に個人差があるため、ある被験者ではボタンが押下してない時でも笑顔の強度が少々出る場合や、笑顔の

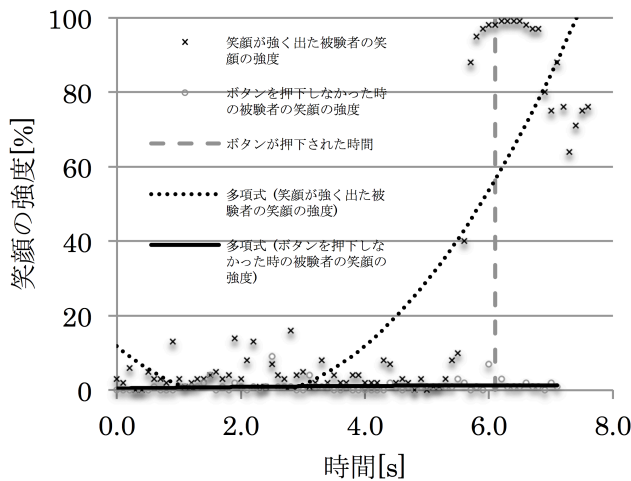


図2 笑顔が強くなった被験者の笑顔の強度

表2 形態素解析を用いた感情極性評価値の算出結果

ID	単語感情 極性対応表	日本語評価極性辞書		面白いと思った 被験者の数
		用言編	名詞編	
1	-0.3578	1	-2	5
9	-0.35458	0	0	4
11	-0.61965	0	0	0
15	-0.44862	1	3	4
31	-0.41799	1	4	0

強度のばらつきがあった。しかし、ボタンが押下した時では笑顔の強度がばらつくことなく強く出ているため、笑顔の強度を測定する被験者固有のしきい値を設定することで感情の変化を検出することができるのではないかと考える。

また、形態素解析を用いての感情極性判別の結果で、日本語評価極性辞書での結果ではIDが11のツイートはニュートラル、15のツイートはポジティブという評価値がそれぞれ出ているが、この結果と同じように被験者から面白いというポジティブな評価が出ている。しかし、IDが1、9のツイートでは評価値としてニュートラル、またはネガティブな評価が結果として出ているが、被験者からの評価ではポジティブな評価が出ている。また、IDが31のツイートでは評価値が高くポジティブという結果が出ているが、被験者からの評価ではポジティブな評価は出していない。これらのことから、ツイートの各単語の感情極性からだけでは感情を推測することは難しいが、本研究の手法を適応し、ユーザの表情から感情を推測することでこの問題を解決できるのではないかと考える。

6. まとめ

本研究ではリツイート時の感情変化に起こる表情の変化に着目し、ツイートしたユーザのセンチメント分析を行った。その結果、感情の変化が起こり、表情に出てからリツイートなどのボタンを押下しているの、常に表情を取得しなくともボタン押下時の表情を取得することで表情からセンチメント情報を取り出せることがわかった。また、ツイートの各単語の感情極性だけから感情を推測することは難しい場合があるが、本研究の手法を適応し、ユーザの表情から感情を推測することでセンチメント分析できるのではと考えた。

今後はポジティブ、ネガティブの2種類だけでなく、他の感情に対しても本手法を適用することを考えている。

参考文献

- [1] "Twitter", <https://twitter.com/> (2013-10-22 参照)
- [2] "ReKognition API", <http://rekognition.com/> (2013-10-22 参照)
- [3] 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.02 pp. 627-637, 2006.
- [4] 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, pp.203-222, 2005.
- [5] 東山昌彦, 乾健太郎, 松本裕治, 述語の選択選好性に着目した名詞評価極性の獲得, 言語処理学会第14回年次大会論文集, pp.584-587, 2008.
- [6] "MeCab: Yet Another Part-of-Speech and Morphological Analyzer", <http://mecab.googlecode.com/> (2014-01-10 参照)