

# Twitter上に投稿された文章に基づく感情推定法と その応用に関する検討

†松林 圭 †五味 京祐 †古川 和祈 †松尾 祐佳 †松原 良和

‡日諸 マルセロ優次 ‡中村 拓哉 †山下 晃弘 †松林 勝志

†東京工業高等専門学校 情報工学科 ‡株式会社調和技研

## 1 はじめに

SNSは誰でも手軽に利用でき、若年者層を中心に広く普及している。SNS上への投稿データや友人関係のリンクに関する研究も数多く取り組まれており[1]、その成果を生かしたマーケティングツールなどの実用化も広がりつつある。一方で、SNSの利用者数の増加と共に「炎上」と呼ばれる投稿者の意図しない投稿の広がりや、それに伴う投稿者への誹謗中傷も増加している[2]。Twitterは代表的なSNSであるが、特にTwitterを発端とした炎上は多い。SNSに投稿される文章には少なからず投稿者の感情が含まれる場合が多く、投稿に含まれる感情を分析することは、文章の構造や内容の解析とあわせて、その投稿がその後に及ぼす影響についてより詳細に分析できる可能性がある。

本研究では、Twitterを対象とし、Word2Vecを用いてツイートのテキストを特徴ベクトルに変換し、ランダムフォレストを用いた感情の分類器を構築することで、ツイートから感情を推定する手法を提案する。提案手法によって得られた分類結果についてはテストデータによる検証の他、アンケートによる妥当性の検証も実施した。また、提案する手法を応用した事例として、Twitterモニタリングシステムへの応用例についても紹介する。

## 2 感情の推定

### 2.1 感情の種類

まず、感情の種類を定義する必要がある。感情表現辞典[3]では、感情を表現する単語を、喜、好、昂、怒、哀、厭、驚、怖、安、の10種類に分類している。しかし、細分化された感情推定は直感的ではなく、本研究の目的であるツイートの感情分類には適さないと判断し、今回はより簡略化して喜、怒、哀、楽、と「無感情」の5つ

### An Emotion Estimation Method from Twitter Tweets and the Application

†Department of Computer Science, National Institute of Technology, Tokyo College.

‡Chowagiken Corp.

に分類することにした。喜怒哀楽はそれぞれ感情表現辞書から喜={喜,好,昂},怒={怒},哀={哀,厭,驚,怖},楽={安}として、感情表現辞書の単語を、それぞれの感情を表現する語として割り当てた。また、本研究では、Word2Vec[6]を用いて感情表現辞書に記載された単語の類義語を検索し、その類義語も喜怒哀楽を表現する単語として追加した。本稿ではこの単語を感情語と呼ぶ。

実験前の事前準備として、一般的に、喜怒哀楽の各分類の感情語がTwitter上にどの程度出現しているか、その割合を調べた。結果を表1に示す。

表1: 感情語がツイートに出現する頻度

	喜	怒	哀	楽
感情表現辞書の単語数	758	207	1078	117
感情語の単語数	20299	4428	26242	2788
ツイートの出現割合 [%]	0.31	0.024	0.25	0.0066

### 2.2 ツイートの感情推定法

感情推定のために、ツイートの特徴ベクトルを生成する。特徴ベクトルの生成方法には様々な手法があるが、単に文章中に感情語が出現する回数を数える方法や、TF/IDF法では、否定文など文の構造を考慮した表現に対応できないうえ、新出単語に対応できない問題がある。そこで本研究では、ツイートを入力として各単語のベクトルを出力するWord2Vec[6]を用いた。Word2Vecで生成されるベクトルは、その距離が近ければ、単語の意味も近いと推定可能である。また、否定文や疑問文になる場合も特徴ベクトルの変化によって分類が可能である。

次に、生成した特徴ベクトルを感情で分類するための分類器を構築する。今回は高速で精度が良く、大量データに適したランダムフォレストを用いて分類を行った。また、Word2Vecの学習にはWikipediaのコーパスを使用した。

## 2.3 感情推定法の評価実験と結果

学習データは、「怒」や、「悲」などの典型的な感情語で検索して得た約 4000 のツイートを目視で確認し、それぞれに喜、怒、哀、楽、無感情のラベル付けを行ったものを使用する。データ数はそれぞれ、喜が 659、怒が 1605、哀が 480、楽が 344、無感情が 700、用意した。テストデータのうち、90%を教師用、10%を評価用とした。

テストデータと評価データを入れ替えて 10 回実験を行った際の正解率の平均値は 74.63%であった。また、感情語で検索したツイートではなく、Twitter 全体からランダムに抽出したツイートに対しても感情推定を実施した。無作為に抽出した 355 個のツイートに対して提案手法で感情推定した結果について、20 代男性 4 名に目視でその結果が、妥当、不適切、どちらともいえない、の 3 段階で評価してもらった。その結果、妥当 35.5%、不適切 40.8%、どちらともいえない 23.7%という結果となった。不適切が多くなった原因は、無感情のツイートに感情ありのラベルを付与してしまった誤りが多いためであり、今後は無感情のツイートの検出精度向上が課題である。また、同じ文章でも、前の文を受けて感情が変わる文章は正しく推定できないため、前後のツイートの感情の推移を考慮した推定方法の検討などが必要である。

## 3 Twitter モニタリングシステムへの適用

筆者らは、Twitter の安心安全な利用促進を目的とし、個人や企業向けモニタリングシステムの開発に取り組んでいる [4][5]。炎上対策はツイートの投稿より 2 時間以内が重要 [2] であり、筆者らが開発を進めるモニタリングシステムでは、可能な限りリアルタイムにツイートの内容、リツイート数、お気に入り登録された数を取得し、一定の伸び率を超えた際に通知する機能などを備えている。また、NG ワードでのフィルタリングなど、不適切な投稿を防止する機能を実装している。

一般的に、炎上等は、その返信に否定的な内容が多く出現する。そこで、本研究で提案した感情推定手法をモニタリングシステムに搭載し、ツイートに対する返信を監視し、感情が否定的な返信が多い場合に通知する機能として実装した。

また、炎上等に関わらず平時においてもツイートの感情の推移を可視化できる機能を実装したほか、個人ユーザへのモニタリングシステムの利用推進のアプローチとして、芸能人やスポーツ選手とユーザーとの感情的な相性判断や、感情の未来予測など、エンタテインメント性を高めた付加機能として実装した。図 1 に開発したシステムの画面を示す。

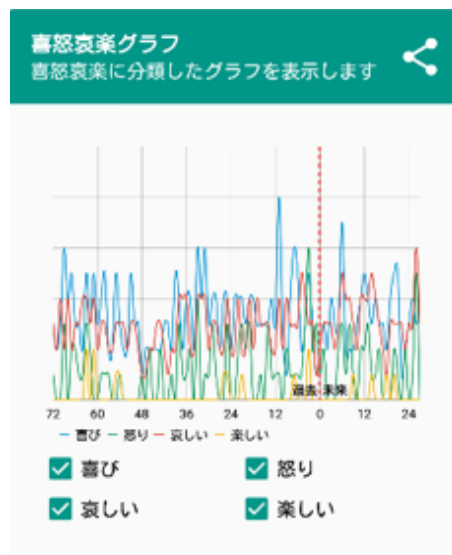


図 1: Twitter モニタリングシステムへの応用例

## 4 まとめ

本研究は、Twitter に投稿されたツイートについて、喜、怒、哀、楽、無感情の 5 つのカテゴリで感情推定手法を提案し、評価実験を実施した。感情推定は、感情表現辞書に基づいて予めラベルを付与したツイートを教師データとし、Word2Vec を用いた特徴ベクトルの生成と、ランダムフォレストによる分類器によって構築した。評価実験の結果、テストデータは 74.63%の推定精度が得られたものの、一般的なツイートを対象とした分類結果では 35.5%の精度に留まった。これは、無感情の推定精度が低いことが影響しているため改善が必要である。また、本手法を Twitter モニタリングシステムに応用した事例を紹介した。今後は学習データを増やして精度を高めるとともに、ツイートの前後関係から推定する方法を検討する。

## 謝辞

本研究の一部は、JSPS 科研費 15K16092 の助成を受けたものです。

## 参考文献

- [1] 大向一輝, 松尾豊, 他: 特集 Twitter とソーシャルメディア, 人工知能学会誌, Vol.27, No.1, pp.34 - 67 (2012).
- [2] 山下晃弘, 他: SNS プライバシー保護とリスク管理の検討, 情報処理学会デジタルプラクティス, Vol.6, No.2, pp.150-158 (2015).
- [3] 中村明: 感情表現辞書, 東京堂出版 (1993).
- [4] SNS 炎上報知器, <http://見守り.jp>, 東京高専.
- [5] Cikappo, <http://cikappo.weboss.co.jp/>, (株) 調和技研.
- [6] Word2Vec, <https://code.google.com/p/word2vec>.