

# ツイートにおける周辺単語の感情極性値を用いた新語の感情推定

黒澤 瞭佑<sup>†</sup> 酒井 哲也<sup>†</sup>

<sup>†</sup> 早稲田大学基幹理工学部情報理工学科 〒169-8555 東京都新宿区大久保 3-4-1

E-mail: †kuro-one@akane.waseda.jp, ††tetsuyasakai@acm.org

あらまし SNS のひとつである Twitter におけるツイートにはユーザの感情が現れやすいため、ユーザを細かくプロファイリングすることを目的としたツイートの感情推定を行う研究が数多く存在する。しかし、従来研究において新語・未知語を含むツイートの感情推定および新語・未知語自体の感情推定を行っているものは少ない。従来研究のひとつに任意の単語に対して語彙ネットワークを利用して自動的に感情極性値を付与した研究があり、その結果を利用して新語の感情極性がポジティブであるかネガティブであるかの判定およびそれに準ずる感情の付与を実現できると考えられる。本研究では、ツイートに含まれる新語の前後に存在する周辺単語の感情極性値には新語の感情極性値に対して相関があるという仮定のもと、新語に対しての感情値の付与および感情推定を行う手法を提案し、その評価を行う。

キーワード Twitter, 新語・未知語, 感情推定, 感情語辞書

## 1. はじめに

スマートフォンを始めとする携帯端末の普及により、人々がいつでもインターネット上でコミュニケーションをとれる時代が到来した。身近なコミュニケーションツールとして提供されている SNS (Social Networking Service) には多種多様なものが存在し、数多くのユーザが発信者・受信者となり多くの情報をやり取りしている。

SNS の一つである Twitter では、ユーザがツイートと呼ばれる 140 文字のショートテキストを公開することで、不特定多数のユーザや、知人・友人間でのコミュニケーションが行われている。また、お気に入りのユーザや知人のアカウントをフォローすることで、そのユーザの公開したツイートを閲覧することができ、個人の嗜好や交友関係に合ったものだけをまとめて表示できるようになっている。ツイートには日常で起こったことやユーザの感じたこと、体験したことなどを主とした、様々な情報が含まれている。

ツイートから得られるユーザに関する様々な情報を分析することで、ユーザエクスペリエンスの向上やユーザへの適切かつ効果的な情報推薦などを行うことができる。実際にユーザが入力したツイートの検索クエリや、フォローしているユーザの類似性、閲覧したツイートの特徴、ツイートの内容、ツイートをを行った位置情報などから、適切な広告の表示や類似したユーザの推薦などが行われている。ユーザの日常に密接しているツイートには、ユーザの感情が多く反映されていると言うことができる。その中でもツイートの感情を推定する研究は今までも盛んに行われており、ユーザのプロファイリングや適切な情報推薦のために利用されている。

ツイートの感情解析をすることで、ユーザのリアルタイムの感情やムードを得ることができ、それにユーザに対する情報の推薦や、ネガティブなツイートを繰り返しているユーザに対してポジティブになれるような提案や自殺などの行動の抑止に

つながると考える。

しかし、Twitter には新語・未知語の出現回数が多く、140 文字までという文字制限による文法の乱れなどから、高精度での感情推定が困難である。松林ら [1] の研究では、一般的なツイートにおけるテストユーザによる評価実験で 35.5% の精度に留まっている。また、村石ら [3] の研究においても、感情推定の精度は約 20.83% となっている。いずれにおいても、ツイートに含まれる単語が参考対象とした感情語辞書に登録されていないことや、辞書自体がツイートのような短文を想定して作成されていないこと、感情値の付与されていない新語・未知語に対して対応できないことなどを問題としている。

本研究では、新語・未知語を含む文章の感情推定精度の向上のために、新語・未知語の感情推定およびその評価を行うことを目的とする。形態素解析ツールである MeCab [10] でツイートを形態素解析し、得られた新語・未知語の感情解析を行う。通常の MeCab 辞書では Twitter に現れる新語・未知語を形態素解析することが困難であるので、mecab-ipadic-NEologd [11] と呼ばれる MeCab 専用のカスタムシステム辞書を用いて感情の付与されていない新語・未知語の検出を行う。得られた単語の周辺に存在する単語を周辺語とし、それらの感情は得られた単語の感情に対して相関を持つという仮説のもと、高村ら [5] の研究で得られた単語感情極性対応表を用いて新語・未知語の周辺に存在する語の感情スコアを算出し、新語・未知語の感情極性をポジティブ/ネガティブの二値で推定する。その後、ユーザ評価を行い、作成された正解データと推定データを比較・評価する。

## 2. 関連研究

松林ら [1] の研究ではツイートデータの解析に Word2Vec [12] を用いて特徴ベクトルを抽出し、ランダムフォレストを用いた感情の分類器を用いることで喜怒哀楽に加え無感情の 5 カテゴリーにわたる感情推定を行っている。堀宮ら [2] の研究では、人間

の人間に対する推測能力に着目してユーザのツイートに対する他ユーザの反応であるリプライを利用して感情推定を行っている。村石ら [3] の研究では、ツイート中からユーザの感情を抽出し感情を「悲しい ⇄ 楽しい」、「安らぎ ⇄ 怖い」、「リラックス ⇄ 怒り」、「高揚 ⇄ 落ち込み」の 4 軸で推定し、抽出された感情と対になる感情を推定し、それに基づく行動推薦を行うシステムの研究を行った。山本ら [4] の研究ではツイート内の特有表現として顔文字を挙げ、顔文字に着目したツイートの感情推定を行っている。顔文字にはツイートの感情を強調、弛緩、転換するなどの効果があり、感情推定に感情語辞書だけでなく顔文字辞書の作成も行っている。

いずれの研究においても、それぞれのアプローチでツイートの感情推定を行っているが、未知語や特有表現、文字数の制約による文法の乱れなどから高い精度を出すことができていないのが現状である。そこで、本研究で新たなアプローチとして、ツイート中に含まれる新語・未知語およびその周辺語に着目して感情推定を行い、得られた新語・未知語の感情極性を辞書化することで、以降の感情推定の精度向上を目指す。本研究の目的は、ツイートに含まれる新語・未知語を抽出し、その周辺単語の感情値から感情を推定し、その精度を評価することである。

### 3. 提案手法

提案手法の概要について図 1 に示した概要図を用いて説明する。まず、Twitter より得られたツイートデータを形態素解析ツール MeCab で品詞分解し、感情語辞書によって与えられた感情極性値を持たない新語・未知語と感情語辞書に登録があり、感情極性値を持つ周辺語を抽出する。感情極性値を持たない新語・未知語にはその周辺語の感情極性値を利用し、それらに対する感情極性値を算出してスコアの付与を行い、感情の分類を行う。その後、提案手法によって付与された感情値をテストユーザが手動でタグ付けした正解データと比較し、適合率を評価する。

以下の項では、評価までの各パートの具体的な内容に関して記述していく。

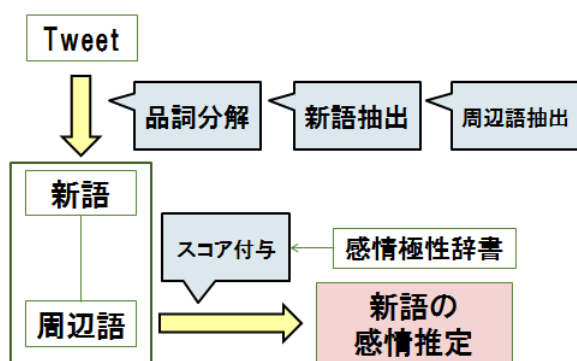


図 1 システムの概要図

#### 3.1 感情語辞書

提案手法を実現する上で必要となる感情語辞書についての説明を行う。感情語辞書とは感情を表す単語を集めた辞書のこと

であるが、一般的には一つの単語に対して一つの感情極性値がセットとして与えられたものの集合を指す。品詞分解を行ったうえでの感情推定においては、単語に与えられた感情値の影響が大きいため、使用する感情語辞書によって実験結果は大きく左右される。感情語辞書の構築は、人手による手動タグ付けで構成されるものと、分類器や深層学習などで自動的に感情極性値を算出し割り振ったもので構成されるものの 2 種に大別される。人手で感情極性を付与した感情語辞書には小林ら [6]、東山ら [7] の研究で得られた日本語評価極性辞書が挙げられる。日本語評価極性辞書には約 5,000 語の用言と約 8,500 語の名詞および複合名詞に対しネガティブ/ポジティブの極性と主観/客観のタグが付与されたものが収録されている。自動で感情極性値を付与した感情語辞書には高村ら [5] の研究で得られた単語感情極性対応表がある。単語感情極性対応表には合計して約 55,000 語の名詞、動詞、副詞、形容詞に対し、語彙ネットワークを用いて算出された  $-1$  から  $+1$  までの実数値が感情極性値として付与されている。この場合、値が  $-1$  に近い程ネガティブで  $+1$  に近いほどポジティブな単語であると判断される。本研究では、主に高村ら [5] による単語感情極性対応表の単語およびその感情極性値を用いて研究を進めていく。

#### 3.2 感情推定方法

感情を推定する方法には、抽出された単語や文章そのものを、堀宮ら [2] の研究の様に「幸福感」「驚き」「恐れ」「悲しみ」「怒り」「嫌悪」などの細かな感情に分類し、感情推定を行うものと、中道ら [8] の研究の様に「ネガティブ」「ポジティブ」に分類して推定するものが存在する。いずれの研究においてもそれらの感情群に、どの感情にも該当しない場合の「ニュートラル」もしくは「無感情」のタグを加えた分類および推定を行っている。本研究では、細分化された感情ではなく、「ネガティブ」「ポジティブ」「ニュートラル」の 3 値での分類および推定を行うものとする。

#### 3.3 単語抽出

ツイートからの単語抽出には形態素解析ツールである MeCab [10] を使用する。MeCab の標準搭載されている辞書には Twitter に次々と現れる新語・未知語が登録されていないため、形態素解析を行った際に不適切な品詞分解が行われ、不適切な結果が出力されてしまう。それを防ぐため、MeCab のカスタム辞書である mecab-ipadic-NEologd [11] を利用する。mecab-ipadic-NEologd は製作者が現在もなお 2 週間に 1 度程度、Twitter に出現した新語・未知語の登録を行っているため、得られたツイートを適切に品詞分解することが可能である。その後品詞分解をして得られた単語をツイートごとにリスト化し、感情語辞書に登録されている単語とのマッチングを行う。感情語辞書にヒットしない単語が出てきた場合それらの単語を新語・未知語とし、それらの前後にある単語を新語・未知語に対する周辺語とする。周辺語には感情語辞書に登録されている感情極性値を付与し、新語・未知語の感情推定をする際に利用する。本研究では、単語の周辺語にはその単語の感情極性に対して相関があると仮定し、周辺語の感情極性値から新語・未知語の感情極性値を推定する。なお、利用す

る mecab-ipadic-NEologd のバージョンは 2016 年 11 月 17 日に更新された mecab-ipadic-2.7.0-20070801-neologd-20161117 とする。

### 3.4 評価方法

3.3 節で得られた新語・未知語に対し、周辺語を用いた感情値付与を行う。新語・未知語の前後に存在する周辺語を参照し、単語感情極性対応表から周辺語に与えられた感情極性値の相加平均を算出する。得られた単語の感情極性値はあらかじめ設定した閾値に従って「ネガティブ」「ポジティブ」「ニュートラル」の 3 値に分類する。前後に感情値がなく、相加平均を算出できない場合には単語の感情値を 0 とし、自動的に「ニュートラル」に分類する。評価対象は、ランダムに選ばれた 100 件のツイート中に含まれる新語・未知語とする。動詞、形容詞、副詞などの用言の新語・未知語は、活用形が多く、適切な感情値付与が困難であることに加えて、前後の文脈や周辺語によって感情の捉え方が大きく変化してしまうため、常に一定の感情値を与えることができない。そのため、評価の対象とする品詞は文脈によって感情の捉え方が大きく変化せず、一定の感情値を与えることができる体言である名詞のみとする。なお、周辺語は用言、体言の両方を対象とし利用する。

取得したツイートの例を表 1 に示す。ツイートの取得には Python 用 TwitterAPI 利用ライブラリ Tweepy [13] を用いた。表中の下線は MeCab と感情語辞書を用いて新語・未知語と判定された名詞を示している。

抽出された名詞をユーザに評価してもらうためのユーザ実験の一例を図 2 に示す。正解データは、評価対象としたランダムな 100 件のツイート中の新語・未知語と判定された名詞に対し、テストユーザに手で「ネガティブ」「ポジティブ」「ニュートラル」の 3 値をタグ付けしてもらうことで作成する。最終的な評価はシステムによる新語・未知語の感情値の評価とテストユーザによる正解データで比較をし、その適合率で行う。

表 1 取得ツイートの一例

学生の論文のドラフト、たまに文字がダミーで「あ」で埋められているのがあって、「また、以下の考察を行った。ああああああああああああああああ」とか書いてあってちょっとホラー
論文の錬金術師
スカやレゲエを聴いて電車で移動していると、気づいたらめっちゃ自分がノリノリになっててヤバイな。
貴方の指の魔剤？
恋人に髪切ってもらい帰りであろう親と落ち合うからムサコのフードコートにいるけど、適度なざわめきでいい場所○
帰宅なう。寒かった。ガンダムはニュースの後で見ると、こないだアイロンで襟を焦がしてしまったシャツの襟を取って、ノーカラーシャツにリメイクした(ω) 悪くない(ω)
テストの日程が被ってはいけない日に被ってオワオワリなんだが

## 4. 評価実験

### 4.1 実験結果

#### 4.1.1 システムによる新語・未知語の感情スコア

新語・未知語の感情は周辺語の感情に依存しているという仮

	A	B	C
1	新語・未知語	極性評価	
2	ドラフト	ニュートラル	
3	考察	ニュートラル	
4	ホラー	ネガティブ	
5	錬金術師	ニュートラル	
6	スカ	ポジティブ	
7	自分	ニュートラル	
8	ノリノリ	ポジティブ	
9	貴方	ニュートラル	
10	魔剤		
11	ムサコ	ネガティブ	
12	フードコート	ニュートラル	
13	適度		
14	ざわめき		
15	ガンダム		
16	ニュース		

図 2 ユーザ実験の一例

説を検証すべく、システムの作成及びそれによる実験を行った。実験に際して、ツイートデータを自身の TL から 100 件取得し、それらを MeCab により品詞分解した。その後、単語極性対応表に対応する単語には感情値を与え、登録の無い単語に対しては 0.0 の感情値を与えた。なお、新語・未知語は名詞のみとし、品詞分解の際に意図せず得られた不適格な名詞はシステム中であらかじめ取り除いている。

不適格とした単語の種類としては、URL や数値、指示語、ユーザネームなどの名詞や、「ワンマ」+「ン」や「もうこん」+「な」などシステム辞書によって適切に分解できないもの、「好き」「ほんとうに」など名詞でないのに名詞と判定されてしまったものなどが含まれる。

抽出された新語・未知語に対しその前後の単語の感情値の相加平均を算出した後、ネガティブな単語に対して  $-0.5$  以下、ポジティブな単語に対しては  $+0.5$  以上、それ以外をニュートラルとする閾値を設定し、全新語・未知語に感情極性を与えた結果を表 2 に示す。なお、数回にわたって出現した単語に対しては平均スコアを算出した。

表 3 にはベースラインとして、各新語・未知語に対して  $-1.0$  から  $+1.0$  のランダムな値を与え、同一の閾値を用いて感情極性を与えた単語の総数を示した。

表 2 感情極性評価 (閾値 =  $\pm 0.5$ ) [単位: 語]

ポジティブ	ネガティブ	ニュートラル	合計
1	12	151	164

表 3 感情極性評価 (閾値 =  $\pm 0.5$ ) [単位: 語]

ポジティブ	ネガティブ	ニュートラル	合計
30	36	98	164

#### 4.1.2 ユーザ実験結果

本研究におけるシステムの評価指標である取得新語・未知語の感情極性の正解データは、ユーザ実験によって決定した。ユーザ実験では、取得した全 164 語の新語・未知語に対して、共起される感情が前向きな感情である場合「ポジティブ」、後ろ向きな感情である場合「ネガティブ」、そのどちらにも当てはまらないまたは感情が共起されない場合「ニュートラル」と回答してもらった。また、その新語・未知語自体を知らない場合にはあいまいな感情を付与させないために「わからない」と回答してもらった。「わからない」の回答を除き、ユーザの回答が一番多かった感情極性を正解とする。各感情における正解データの数を表 4 に示す。回答票が同数だった場合には「ニュートラル」を除く感情極性で一番多い回答を選択した。なお、ユーザ実験の被験者は 10 名である。被験者に関する情報は付録に記載した。

表 4 各感情における正解データ数

ポジティブ	ネガティブ	ニュートラル	合計
20	21	123	164

#### 4.2 評価

システムによって生成された新語・未知語の感情評価とユーザ実験により得られた正解データとの比較した結果を以下に示す。システムとランダムに値を与えたベースライン Base.1 と全ての単語をニュートラルと判定した Base.2 において全取得新語・未知語の正解データとの適合数および適合率を表 5 に示す。また、システム評価における各感情ごとの適合率を表 6 に示す。

表 5 システムとベースラインの正解データにおける適合数・適合率

	適合数	適合率
システム	114	0.6909
Base.1	80	0.4848
Base.2	123	0.7500

表 6 システム評価における各感情極性ごとの適合率

ポジティブ	ネガティブ	ニュートラル
0.0000	0.0833	0.8750

### 5. 結果・考察

ここでは、得られた結果とその評価を基にした本システムの結果の考察と、その原因についての考察を行う。

表 5 をみると、ランダムに感情値を付与したベースライン Base.1 の適合率が約 49%であるのに対し、提案システムによる適合率は約 69%となっており、ランダムなベースラインよりも提案システムによる評価の適合率の方が高いことが分かる。これは符号検定 [9] による p 値からも明らかである。

$$p\text{-value} = \frac{2 \sum_{k=0}^{n'} n' C_k}{2^{n'}} = \frac{\sum_{k=0}^{n'} n' C_k}{2^{n'} - 1}$$

上式で算出した p 値は  $1.74285E - 05$  となり、「両システムの母集団の分布に差はない」という帰無仮説は棄却され、提案システムの有効性が統計的に優位であることがわかる。

しかし、提案システムによる評価の感情別の適合率（表 6）を見ると、ポジティブ・ネガティブと判定された新語・未知語の正解データとの適合率はそれぞれ 0.0000%, 0.0833%となっている。取得できた全 164 語の新語・未知語の中で、ポジティブな単語が 1 語、ネガティブな単語が 12 語と少ないことを考慮しても、ほとんど正解データに適合しておらず、適合率が非常に低いということが分かる。

また、実験データが極端にニュートラルに偏っているため、全ての単語をニュートラルとする Base.2 の適合率は 75%となり、提案システムはこれに及ばない結果となっている。

このような結果となった原因を、取得単語数、周辺語への対応、文脈の単語への影響、感情閾値の有効性の 4 項目にあるとし、それぞれに対する考察を以下に示す。

#### 5.1 取得単語数

本システムで解析したツイートは 100 件で、その中に含まれていた新語・未知語と判定された単語は重複を含まず数えると 164 語であった。新語・未知語の数としては解析するにあたって十分な量であると考えられるが、周辺語に感情極性値を持つものが少なかったため、平均値を算出することができなかった新語・未知語が多く出てしまった。これを解決するためには、更に多くのツイートを解析し、重複した単語の感情極性値に対しては得られた値の平均値を算出することで改善できると考えられる。

#### 5.2 周辺語への対応

本システムにおいて、新語・未知語の周辺語は必ずしも感情極性値を持っているとは限らないということを考えていないということが挙げられる。各取得新語・未知語に対しての前後の単語の感情値を算出した数値を記録した際、得られた新語・未知語の感情値が 0 と出力されているものが多く見られた。これは、新語・未知語の周辺語に感情値の無い場合には、平均値を算出することができないため、該当単語の感情値を 0 としたからである。つまり、感情極性値の平均が 0 となっている単語に関しては周辺語から感情値を取ることができていないため、感情推定以前の問題であり、システムとしては不十分であると言える。これを踏まえてとることのできる対策は二つあると考える。一つは周辺語が感情極性値を持たない場合はさらにその周辺語を取得するという方法を取ることである。もう一つは各新語・未知語をキーワードとして含むツイートを多数取得してシステムに入力として与え、感情極性値の平均値を算出するという方法である。

#### 5.3 文脈の単語への影響

単語の出現する文脈や文全体の感情によって、同じ新語・未知語でも感情に違いがあるということを考えていないことも原因の一つであると考えられる。例えば、「スカ」という単語は「はずれ」という意味と「音楽のジャンルのひとつ」という意味の 2 つを有する。前者はネガティブな感情を有するが、後者はそのジャンルが好きな人にとってはポジティブ、もしくはニュー

トラルだと捉えられると推測できる。このように、単語だけ見てもこれがどちらの意味であるのかを断定することはできないため、感情推定を行う際には単語を含む分の文脈や全体の感情を考慮しなければならない場合があると言える。

#### 5.4 感情閾値の有効性

本研究では、「ポジティブ」「ネガティブ」「ニュートラル」の感情極性に分類する際、感情極性値にネガティブな単語に対しては  $-0.5$  以下、ポジティブな単語に対しては  $+0.5$  以上、それ以外をニュートラルとする閾値を設定した。これは、単語感情極性対応表においてポジティブ・ネガティブと判断された単語の感情極性値の平均値がそれぞれおよそ  $+0.5$ ,  $-0.5$  であったため設定した。しかし実際に周辺語の感情極性値の平均を算出してみたところ、マイナスの数値となるものが多いことが分かった。これは感情極性辞書として利用した単語極性対応表の辞書作成システムによって単語の感情を数値化した際、人間が実際に感じる感情とシステムによる感情極性値の不一致に依るものと考えられる。その感情の不一致を補正するのが閾値であり、今回の実験ではその閾値が適切でなかったことが原因だと考察できる。よって、閾値をより適切な値に変更することによって、その不一致を解消し、精度を向上させることができると考えられる。

## 6. 追加実験

考察を踏まえてシステムの推定精度を向上させるための追加実験を行った。ここではその追加実験の内容とその結果、および考察を行う。

### 6.1 実験

#### 6.1.1 本実験の意義

5.2 節では、取得した新語・未知語の周辺語に感情極性値が無い場合を考慮し、新語・未知語をキーワードとして検索して得られた多数のツイートの感情極性値の平均値を取ることで、感情値を取得することができると考察した。まだ、5.4 節では、適切な閾値の設定によって本システムを用いた感情推定の精度を向上することができるであろうと述べた。

今回の追加実験は、新語・未知語をキーワードとして得られたツイートから感情極性値の平均を算出し、その結果を踏まえて各感情に対して適切な閾値を設定することができるか調査することを目的とする。

#### 6.1.2 実験内容

正解データにおいて「ポジティブ」「ネガティブ」「ニュートラル」と判定された単語をキーワードとしたツイートを各単語 100 件取得し、作成した評価システムで感情極性値を算出し、その平均値を出力する。なお、キーワードとした単語は、回答票数の多かった「ポジティブ」「ネガティブ」な単語上位 10 単語ずつ、「ニュートラル」は全員がニュートラルと回答した単語のうち、人物、場所、動物などを名前を除く 25 単語とした。実際に対象とした単語は以下の表 7 に示した。

#### 6.1.3 実験結果

各感情における感情極性値の平均を表 8、表 9、表 10 にそれぞれ示す。

表 7 キーワードとした新語・未知語

感情極性	新語・未知語
ポジティブ	神曲, 大人気, キュン, 素敵, サンキュー, キタコレ, きれい, 草, ノリノリ, 結婚式
ネガティブ	ネガティブ, エラー, メンヘラ, ブス, 罪悪感, 無理, ホラー, 墓場, ネタバレ, 卒論
ニュートラル	状態, 出だし, 説, こないだ, 日, 図書館戦争, 人, ツイート, LINE, 年会, 普段, 前, やり方, 感じ, 日本, インスタ, fox, チャリンコ, 気分, 合宿, 考察, TL, 普通, ブログ, 人生

表 8 ポジティブな単語に対する評価

ポジティブ	ave_pos
神曲	-0.0872
大人気	-0.1697
キュン	-0.1058
素敵	-0.2535
サンキュー	-0.0610
キタコレ	-0.0876
きれい	-0.1139
草	-0.0759
ノリノリ	-0.1006
結婚式	-0.1494

表 9 ネガティブな単語に対する評価

ネガティブ	ave_neg
ネガティブ	-0.1946
エラー	-0.1485
メンヘラ	-0.2058
ブス	-0.1376
罪悪感	-0.2232
無理	-0.1334
ホラー	-0.1723
墓場	-0.1584
ネタバレ	-0.1824
卒論	-0.1070

## 7. 考察

得た表を用いて結果の考察を行う。まず、値が 0 となって出力された単語を詳しく調べた。すると、周辺語に感情値を持たず、全く同一の内容のツイートが取得されているものや、単語が中途半端な形で形態素解析されているために当てはまる単語の前後に感情値が無い場合などに感情値が算出できずに 0 になってしまうことが分かった。

次に、追加実験により得られたすべての値は 0 を下回り、マイナスの値となっていることが分かる。これにより、ポジティブな単語であっても値がプラスになるわけではないことが分かり、人間の評価とシステムの評価の不一致が明確となったと言える。

では、この結果から適切な閾値が設定できるか考察する。出力された各値の平均を求めたところ、表 11 となった。同時に各感情の最小値・最大値も記録し、どの各感情の各感情極性値は程度の範囲に及ぶのかをまとめた。

表 10 ニュートラルな単語に対する評価

ニュートラル	ave_neu
状態	-0.2199
出だし	-0.1504
説	-0.1541
こないだ	-0.0852
日	-0.1191
図書館戦争	-0.1591
人	-0.1811
ツイート	-0.1673
LINE	0.0000
年会	0.0000
普段	-0.1942
前	-0.2334
やり方	-0.2250
感じ	-0.0901
日本	-0.1476
インスタ	-0.1416
fox	0.0000
チャリンコ	-0.1432
気分	-0.0932
合宿	-0.1920
考察	-0.1536
TL	0.0000
普通	-0.1296
ブログ	-0.0179
人生	-0.1381

表 11 各感情の平均感情極性値とその範囲

	ポジティブ	ネガティブ	ニュートラル
平均値	-0.1205	-0.1663	-0.1493
最小値	-0.2535	-0.2232	-0.2334
最大値	-0.0610	-0.1070	-0.0179

まず、各感情ごとの平均値を比較すると、ネガティブ < ニュートラル < ポジティブの順に平均値が高くなっていることが分かる。この結果からある程度の相関はあると推測できる。しかし、これらの値の差はとても小さく、各感情を十分に差別化できるとは言えない。また、同様に最小値・最大値で評価値の範囲を比較してみてもそれぞれの差はほとんどないため、各感情を推定するには十分でないと言える。

## 8. 結 論

本研究では、新語・未知語の持つ感情を推定するために、単語の周辺語の感情極性値を用いた推定方法を提案し、実験を行った。最初の提案手法による実験では精度よく推定することができなかったが、追加実験では新語・未知語の周辺語の感情極性値にはある程度の相関があることが分かった。各感情はそれぞれ個別の感情極性値の範囲を持っているが、範囲が重複している箇所があることが分かった。故に、今回の実験のように一定の閾値を与えることによって行う感情分類は難しいと言える。

## 9. 今後の展望

今回の研究では、新語・未知語はどのツイートでも必ず同じ感情極性値を持つと仮定して感情推定を行ったが、単語の感情極性値はツイート毎にある程度の振れ幅があるがあるということをも前提にして推定を行うと精度を上げることができるだろうと考えた。そのために、ツイート自体の全体的な感情を分析し、そこから得られた新語・未知語がそのツイート自体の感情とどの程度相関があるかを研究する必要があるように感じた。

新語・未知語の感情推定手法の研究は、まだどの研究においても高い精度を出すことができていないため、今後も積極的に行われていくべき研究分野のひとつであると考えられる。

## 文 献

- [1] 松林圭, 五味京祐, 古川和折, 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, “Twitter 上に投稿された文章に基づく感情推定手法とその応用に関する検討,” 情報処理学会第 78 回全国大会, 2016.
- [2] 堀宮ありさ, 板野遼平, 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.
- [3] 村石将嗣, 北山大輔, “ツイートの感情推定に基づく対になる感情を誘発する行動の推薦手法,” DEIM Forum 2016, P2-2, 2016.
- [4] 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.
- [5] 高村大也, 乾考司, 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 Vol. 47 No. 2, pp.627-637, 2006.
- [6] 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.
- [7] 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.
- [8] 中道龍三, 徳久雅人, 村上仁一, 池原悟, “情緒推定の手がかりとなる接続表現の収集,” 電子情報通信学会技術研究報告書. TL, 思考と言語 108(353), pp.1-6, 2008.
- [9] 酒井哲也, “情報アクセス評価方法論 検索エンジン評価のために,” コロナ社, pp.152-154, 2015.
- [10] “MeCab: Yet Another Part-of-Speech and Morphological Analyzer,” <http://taku910.github.io/mecab/>.
- [11] “mecab-ipadic-NEologd: Neologism dictionary for MeCab,” <https://github.com/neologd/mecab-ipadic-neologd/blob/master/README.ja.md>.
- [12] “Word2Vec,” <https://deeplearning4j.org/ja/ja-word2vec>.
- [13] “Tweepy,” <http://www.tweepy.org/>.

## 付 録

ユーザ実験に協力していただいた 10 名の被験者の情報を以下に記載する.

表 A.1 ユーザ実験被験者情報

ユーザ	年代	性別	職業
1	20代	男性	学生
2	20代	男性	学生
3	20代	男性	学生
4	20代	男性	学生
5	20代	男性	学生
6	20代	男性	学生
7	20代	男性	学生
8	20代	男性	学生
9	20代	女性	学生
10	10代	女性	高校生