

## 感情分類を用いた単語分散表現からの感性情報抽出 Extraction of Emotional Information from Distributed Representation Using Emotion Classification

松永 拓己<sup>†</sup> 松本 和幸<sup>†</sup> 吉田 稔<sup>†</sup> 北 研二<sup>†</sup>  
Takumi Matsunaga Kazuyuki Matsumoto Minoru Yoshida Kenji Kita

### 1. はじめに

近年、自然言語処理の分野では単語の一般的ではない用法の検出[1]など、単語分散表現から意味推定を行う技術が研究されている。それらの研究では分散表現全体から単語の情報を抽出しており、感性以外の情報が多く含まれている。

本研究では感性情報のみを抽出することでその精度上昇を目指す。その手法として単語分散表現から感情分類を行うニューラルネットワークを作成し、ネットワークが学習した重みから感性情報が強く表れている次元の情報を抽出する。評価方法としては、コサイン類似度を用いて抽出次元同士が類似した単語の比較を行う。

### 2. 関連研究

学習済みの単語分散表現から情報を抽出する研究には、類似する単語の分散表現が類似するという性質を利用し、類似した単語に類似した離散符号を割り当て、特徴を保持したまま記憶容量の削減を行う手法 [2]や、単語をカテゴリで分類しその重心ベクトルを用いてベクトルの要素が特定の次元に集中しているかについての調査 [3]などが行われている。

[2]では単語埋め込み行列において情報を十分に含む離散符号を獲得できたと結論づけており、[3]においても各カテゴリの意味を表現する特有の次元が存在していることが示された。

また、Wang ら[4]は、感情の埋め込みを行う手法である Emo2Vec を提案している。Emo2Vec は、レビュー文からの意見分析を目的として、ローカルとグローバルの2つの異なるモデルに基づき、感性情報を単語ベクトルに追加している。この方法は、Plutchik の感情の輪に基づき、マルチタスク学習を用いることで、既存の感情極性よりも高い表現力を実現している。彼らの研究では、中国語や英語を対象とした実験において、既存の単語埋め込みや感情埋め込みの精度を改善している。

単語分散表現をもとに、感情の埋め込みを学習する手法は既に研究されているが、日本語の単語に対し、感情極性ではなく複数の感情カテゴリを割り当てるための手法はまだほとんど存在しない。そのため、単語分散表現からの感性情報の抽出を実現できれば、単語の感情を表現し分散表現獲得が可能となり、日本語文の感情分析に役立てることができると考える。

### 3. 提案手法

本節では、本研究で用いた感情分類や手法に関する概要を記述する。

### 3.1 本研究で用いた感情

本研究では、教師あり学習を行い単語分散表現から感情分類を可能とするニューラルネットワークを作成し、そこから感性情報の抽出をすることを目的としている。そのため、まず単語の意味がポジティブであるかネガティブであるかの2値分類を考え、感情を強く有していない単語が想定されることからニュートラルを加えた3種類の感情極性を考えた。

また、ポジティブとネガティブにおいて、感情極性が真逆となるよう対の感情を含むことで、それらの感情の情報が関連づけられて記録されているかをニューラルネットの学習結果から判断できると考えた。そこで、Paul Ekman[5]が提唱している6種類の感情を参考とし、ポジティブな感情を「好、褒、驚 P、喜、望」とし、同様にネガティブな感情として「嫌、怒、驚 N、悲、恐」と設定した。また、今回の実験において「好」は「嫌」の、「褒」は「怒」の逆の意味となるように定義した感情である。

### 3.2 感性情報の抽出手法

本研究では  $n$  次元の分散表現を入力とし、全結合層を経由して、入力した単語が前節で定義した3種類の感情極性及び10種類の感情のそれぞれである確率を出力とするニューラルネットワークを構築し、これを学習して感情分類モデルを構築する。そして、以下で述べる2種類の手法を用いてネットワークが感情の分類において注目している次元の調査を行った。

#### 3.2.1 ネットワークの重みを用いる手法

ニューラルネットワークにおける入力の  $n$  次元それぞれに対して重みを計算し、その重みの絶対値の大きな次元が感性情報を多く含む次元であると判断する。このようにする理由は、ネットワークの重みが感情分類をする上での判断根拠となっていることから、ターゲットとなる感情の情報が多く含まれていると考えられるからである。

たとえば、全結合層が1つのネットワークの場合、 $i$  次元目の感情  $e$  に関する重みは  $i$  次元目の入力から感情  $e$  への重み  $w_{i,e}$  となり、全結合層が2つあるネットワークの重みの場合に、2つ目の全結合層の次元数を  $o$ 、次元の番号を  $j$  とすると、式(1)のように表すことができる。

$$w_{i,e} = \sum_{j=0}^o w_{i,j} w_{e,j} \quad (1)$$

さらに全結合層を  $X$  個とした場合、式(2)のようになる。

$$w_{i,e} = \sum_{j=0}^o w_{1,i,j} \sum_{k=0}^p w_{2,j,k} \sum_{l=0}^q w_{3,k,l} \dots \sum_{x=0}^y w_{X,x-1,x} w_{e,x} \quad (2)$$

ただし  $o, p, y$  は中間層の次元数である。

<sup>†</sup> 徳島大学 Tokushima University

### 3.2.2 ネットワークの予測結果を用いる手法

前項で述べた手法では、ネットワークの重みの大きさを計算することで、感性情報が含まれる次元を選択した。一方、ここでは、ニューラルネットが推測する「特定の感情である確率」を比較することにより、その感情となる可能性が高い次元を選択する手法を提案する。この次元ごとの情報は、入力に目的とする  $i$ 次元目のみが 1 で他の次元の値が全て 0 となる One hot ベクトルを入力する事で得られる。

前項の重みを用いる手法では、重みの絶対値の正負が与える影響を考慮できない問題が存在していた。しかし、この手法で得られる予測結果の各次元の値は確率であり、符号を考慮する必要はなく、単純に大きさを比較できるという利点がある。

## 3.3 評価

### 3.3.1 評価データの作成

前節で述べた手法を用いることで、感性情報を多く含む次元を選択することが可能であると考えられる。しかし、注目されている次元のみを取り出すだけでは、その次元が真に感性情報を含むのかを判別することができない。

そこで本研究では、ネットワークの学習に用いた分散表現から感性情報を多く含むと予測された複数の次元を抽出し、抽出した次元の値を用いてコサイン類似度を計算することで、元の単語分散表現を用いた場合と比較する。これにより、適切に感性情報を抽出できているかを評価する。

特定の次元のみを抽出して得られるベクトルは、注目している部分での一致度が高くなるため、類似度が上昇することが考えられる。しかし、元の単語分散表現と比較してコサイン類似度が上昇した単語が存在したとして、その理由が、共通する情報が含まれていたからであるのか、それとも部分的に偶然一致するように学習されていたかであることを判別することができない。

そのため、本研究では元の分散表現を用いた場合において、対象とする単語と類似度の高い単語集合から、上位 100 語を抽出し、それらに対して、抽出した次元のみを用いて対象単語とのコサイン類似度を求めることにした。これによって注目次元のみが偶然類似するようなベクトルを持つ単語が上位に出現してしまう問題を回避できる。

### 3.3.2 評価手法

3.2 節で提案した手法が抽出する次元は、目的とするそれぞれの感情を強く有していると考えられる次元であり、そのコサイン類似度を計算することによって得られる値は感性が強く表現される次元における類似度となる。

たとえば、単語「暑い」と単語「涼しい」の間の、感情「好」に関する類似度を考える。このとき、どちらも温度に関する単語であるが「暑い」は一般に、ネガティブな意味で用いられ、「涼しい」はポジティブな意味で捉えられることが多いと考えられる。このことより、感情「好」に関して、これらの単語間の類似度が低くなると予想できる。

同様に、特定の感情を表すような単語の場合、該当する感情に対し、元の分散表現の類似度よりも、提案手法による感性情報を含む次元のみを用いたベクトル間の類似度が高くなることが予想される。そのため、本研究では評価用の単語と、感情ラベルの教師有学習に使用された単語（例:

「好き」、「褒める」等）を対象に類似度を調査し、元の分散表現と比較することにより評価を行う。

## 4. 実験

本研究は、既存の手法で学習させた分散表現から感性情報を抽出することが目的である。そのため、本実験では Word2vec[6]を分散表現の学習に用いた。また、ニューラルネットワークの実装においては Keras を使用した。

### 4.1 実験準備

#### 4.1.1 実験データの準備

本実験で主に用いたデータは Twitter API を使用しツイートを収集したものである。収集したツイートコーパスに対して、前処理として日本語形態素解析器 MeCab を利用し、分かち書き処理を行った。このとき、記号などのノイズとなるような語の除去や、単語の活用形を全て原形に統一した。活用形を原形にした理由としては、係り受けによる解析を行わない場合、活用形にかかわらず、得られる感性は等しくなると考えたためである。

この結果、約 1.2GB のコーパスが得られた。このコーパスから Python の gensim モジュールを利用し Word2vec を用いて単語分散表現の学習を行った。ハイパーパラメータとして、学習アルゴリズムを CBOW(Continuous Bag-of-Words)、単語の最低出現回数 min\_count を 10 とした。結果として、得られた単語分散表現の語彙数は 77,437 となった。以下、このデータを学習データとして実験を行う。

また、用意したコーパスの約 1%に当たる出現頻度上位 745 単語に 3 種類の感情極性及び 10 種類の感情のタグを人手により付与し、感情分類の正解データとした。このタグは、それぞれの単語ごとに当てはまる感情極性および感情カテゴリ全てを付与したものである。

たとえば、表 1 の単語「悪い」はネガティブな感情を持っていると考えられたため、「驚」を除く負の感情をすべて付与した。単語「恋愛」は「好・嫌」のいずれかの感情を持つと考えられたため、それら両方のタグを付与した。

また、本研究では 3.1 節で定義した 10 種類の感情を含まない単語のみをニュートラルと考え、たとえば、「水」のような感情を持たないと考えられる単語にのみ「ニュートラル」のタグを付与した。

#### 4.1.2 評価データの作成

学習させた感情分類が本当に正しいものであるかを確認するために、「A 群:感情を有していると考えられる単語 25 語」、「B 群:主に他の単語を修飾する単語 25 語」、「C 群:主に他の単語に修飾される単語 10 語」の計 60 単語に対して被験者数名に対するアンケートを実施した。

表 1 タグ付与例

| 単語 | 付与した感情        |
|----|---------------|
| 悪い | 嫌,怒,悲,恐,ネガティブ |
| 恋愛 | 好,嫌           |
| 寒い | 嫌,ネガティブ       |
| 水  | ニュートラル        |

表 2 アンケートを実施した単語の例

|     |  |
|-----|--|
| A 群 | 魅力, 無茶, 目新しい, エlegant, 遠慮ない, 損じる, 速やか, 二流, 苦しむ, 罰当たり 等 |
| B 群 | 消える, 独立, インテリ, やめる, オリジナル, 成長, 超人, あったかい, 理想, 溺れる 等    |
| C 群 | タイミング, ディズニー, 兄さん, 手紙, 映画, 小学校, 塩分, 運営, ゴールデンウィーク, 味方  |

アンケートを実施した単語の例を表 2 に示す。また、これらの単語はそれぞれの条件の下で出現頻度の高い単語が選択される確率が高くなるようにして、ランダムに選択した。

なお、アンケートの項目は以下のようにになっている。

- ・ 対象の単語がポジティブかネガティブか。
- ・ 対象の単語の感情極性がニュートラルであるか。
- ・ 近い感情を 10 種類から上位 3 件まで答える。
- ・ 対象の単語が感情を持たないと考えられるかどうか。

このアンケートの回答は 20 代の男性 6 名を被験者として行った。このアンケートの結果に基づき、それぞれの単語に対して単語の感情極性のポジティブ/ネガティブの一致を求める「2 値分類」、 「2 値分類」にニュートラルを加えた「3 値分類」、アンケートで求めた単語の感情が予測された感情と一致すれば正解とする「most」、単語の感情が予測の上位 3 位以内に含まれていれば正解とする「like」の 4 つの評価項目を設定した。

なお、アンケート結果から単語の感情を決定する際には感情が選択された順に[4, 2, 1]と重みを付け、最も高い感情と、その 3/4 以上の重みを持つ感情のうち上位 2 つを合わせて最大で 3 つの感情を持っていると仮定する。

## 4.2 実験環境

本節では前節で準備したデータを用いて、学習させた感情分類のニューラルネットワークの評価を行う。

### 4.2.1 ニューラルネットワークの作成・学習

4.1.1 で作成した分散表現の次元数にしがって入力を 100, 200, 400 次元、中間層は全結合層 3 層とそれぞれに 0.2 の確率で Dropout をさせる Dropout 層を組み込み、それぞれの感情となる確率を softmax 関数で正規化して出力する全結合層を 1 つ持つネットワークを構築し、損失関数として交差エントロピー誤差を用いた。中間層の次元数は 100 次元の場合は[64, 40, 26]、200 次元では[128, 64, 32]、400 次元では[256, 128, 64]としている。

また、教師データはアンケートの対象として選択された単語が訓練データとして採用されないようにおおよそ 11:1 になるよう分割して学習を行った。学習は最大で 200 epoch 繰り返すようにし、過学習防止のため損失関数が増加した場合に学習を終了する早期打ち切りを設定した。

### 4.2.2 感情分類の評価

前項のニューラルネットワーク並びに、ネットワークの精度の評価を行った。評価項目は 4.1.1 で定めたものを用い、それぞれ 100 回ずつ試行した結果の正答率の平均値を表 3 に示す。

表 3 感情分類の精度の平均

|      | 100 次元 | 200 次元 | 400 次元 |
|------|--------|--------|--------|
| 二値極性 | 0.70   | 0.72   | 0.73   |
| 三値極性 | 0.42   | 0.45   | 0.46   |
| most | 0.27   | 0.26   | 0.27   |
| like | 0.53   | 0.53   | 0.52   |

この結果より、今回作成したネットワークは分散表現のみを入力としていることもあり、極性判定の精度はあまり高くないことが分かる。これは係り受け関係を用いて極性の反転を考慮していないことや、感情タグ付与の対象とした頻出単語の多くが、助詞や「する」といったニュートラルな単語であったことが原因と考えられる。

一方で、本研究において重要となる感情の分類は最も可能性が高いもので 27% 程、上位 3 件まで見れば 50% を超えており、それぞれの単語と感情との結びつきについて学習できているのではないかと考えられる。

また、今回調査した範囲では、感情推定の精度が次元数と関係なく一定であることから、本実験で用いた学習データに関しては 100 次元の分散表現でもデータが圧縮されずに 400 次元と同等の感性情報を有しているのではないかと考えられる。したがって、以降の次元数の差に関する調査においては、100 次元と 400 次元を採用する。

## 4.3 ネットワークの重みを用いた感情抽出

本節ではネットワークが感情分類を行う判断の根拠とした重みに注目して感性情報の抽出を試みる。

### 4.3.1 重みの算出

4.2.1 項で作成したネットワークには、中間層が 4 層存在しており、3.2.1 項にて定義したように全結合層の重みの積を取ったものとする。

また、ニューラルネットワークの重みは通常、正負が混在しているが、掛け合わされる分散表現の値も正負の符号が混在しているため、各次元の意味を値から推定することは不可能である。そのため、ここでは単純に絶対値が大きなものを抽出対象の次元として採用することとする。

上記のようにして得られた重みの大きな次元を各感情に対して 20 次元ずつ選択し、学習元の分散表現から選択された次元の値を抜き出したものを、感情ごとに抽出した感性情報が含まれるベクトルとする。

### 4.3.2 抽出された次元

表 4 は、100 次元の分散表現においてネットワークが注目した 20 次元について、上段左からそれぞれ絶対値が大きな順に並べたものである。ここで、ポジティブな感情に注目してみる。「褒」を除く感情では[99, 95]の次元が最も大きな重みを有しており、「褒」においてもこの次元が出現していることが分かる。

ネガティブな感情においても「驚 N、恐」を除く感情では[29, 17, 15]の次元が共通して最上位に出現している。ネガティブな感情のうち、「驚 N」において、他の感情とは異なる傾向が見られた。この理由として、学習データの不足が考えられる。「驚 N」は、事例数の不足により「驚 P」と同様の感情として学習されてしまったと考えられ、「恐」においても[29, 15]が出現していることから「恐」の感情に固有の感性情報が重視された結果だと推測できる。また、



「褒・怒」での[14, 9, 8, 7]や, 「驚」の[61, 20], 「望・恐」の[65, 64]などは, 対となる感情で共通していることが分かる。

表4 ネットワークが注目した次元1

|    |  |
|----|--|
| 好  | [99, 95, 89, 62, 51, 44, 36, 17, 11, 9, 8, 5, 4, 10, 19, 16, 13, 20, 18, 15] |
| 嫌  | [29, 23, 17, 16, 15, 11, 9, 7, 3, 2, 19, 18, 14, 12, 5, 6, 10, 21, 27, 26]   |
| 褒  | [19, 17, 14, 11, 9, 8, 7, 5, 4, 13, 16, 15, 12, 18, 21, 99, 93, 88, 95, 94]  |
| 怒  | [29, 25, 17, 15, 14, 11, 10, 9, 8, 7, 3, 2, 28, 22, 20, 19, 16, 13, 5, 21]   |
| 驚P | [99, 95, 81, 65, 61, 55, 20, 16, 15, 14, 11, 8, 6, 17, 7, 12, 13, 10, 9, 5]  |
| 驚N | [99, 61, 32, 25, 20, 14, 2, 5, 4, 1, 6, 46, 23, 16, 15, 11, 31, 29, 26, 22]  |
| 喜  | [99, 95, 72, 71, 43, 17, 12, 11, 8, 7, 4, 13, 16, 14, 10, 9, 5, 3, 2, 1]     |
| 悲  | [29, 25, 17, 15, 14, 11, 9, 7, 3, 2, 16, 12, 8, 5, 20, 19, 13, 10, 6, 21]    |
| 望  | [99, 95, 73, 65, 64, 62, 44, 36, 31, 15, 9, 8, 7, 5, 4, 10, 11, 3, 0, 1]     |
| 恐  | [65, 64, 57, 53, 48, 46, 43, 29, 28, 26, 23, 16, 15, 11, 9, 8, 7, 4, 3, 2]   |

表5は400次元の分散表現から表4と同様に次元を抽出したものである。400次元で学習させた場合では, 前述の100次元のときのように広く共通して現れる次元は存在していない。これは1つの次元が表現する情報量が減少したことや, 全体の次元数が増加したことで抽出する次元が占める割合が減少したことが考えられる。

しかし, 400次元の「好・嫌」及び「褒・怒」においても[270, 224, 221, 202, 181]が共通して現れ[275, 185]も多くで見られる他, 「望・恐」において[297, 200, 194, 160, 130, 32]が共通するなど近い感情は同じ部分に着目していることが分かる。

表5 ネットワークが注目した次元2

|    |  |
|----|--|
| 好  | [323, 280, 275, 270, 238, 224, 202, 199, 185, 183, 182, 181, 158, 123, 107, 60, 42, 29, 28, 21]    |
| 嫌  | [275, 270, 252, 238, 224, 202, 195, 185, 181, 158, 113, 107, 42, 30, 11, 28, 1, 36, 19, 15]        |
| 褒  | [321, 270, 255, 251, 224, 209, 202, 200, 181, 133, 115, 108, 101, 100, 81, 62, 60, 40, 29, 28]     |
| 怒  | [317, 275, 270, 224, 214, 202, 199, 185, 183, 182, 181, 60, 40, 28, 316, 295, 145, 120, 108, 100]  |
| 驚P | [127, 120, 29, 28, 10, 1, 16, 5, 2, 15, 3, 8, 9, 11, 14, 19, 26, 25, 24, 23]                       |
| 驚N | [346, 270, 224, 214, 181, 133, 120, 115, 28, 124, 114, 108, 100, 82, 67, 62, 32, 29, 26, 24]       |
| 喜  | [160, 130, 127, 24, 18, 17, 15, 13, 11, 10, 8, 6, 5, 2, 1, 0, 3, 4, 9, 7]                          |
| 悲  | [313, 304, 301, 290, 286, 282, 268, 245, 244, 227, 221, 209, 204, 200, 191, 183, 171, 127, 97, 85] |
| 望  | [309, 297, 200, 194, 160, 130, 95, 32, 26, 24, 18, 14, 12, 10, 4, 6, 7, 22, 21, 20]                |
| 恐  | [221, 130, 24, 21, 16, 11, 5, 2, 3, 7, 18, 19, 32, 297, 203, 200, 194, 160, 94, 67]                |

#### 4.4 ネットワークの予測結果を用いた感情抽出

本節は3.2.2項で提案したように, 4.2節で作成したネットワークの各次元に One-hot ベクトルを投入し, ネットワ

ークの予測した各感情である確率から抽出する感情を含む次元を決定する。その際, ネットワークの予測値が同じとなりやすいため, 100回試行した際の出現回数を取ることに対応した。

##### 4.4.1 次元ごとの比較

表6 ネットワークが注目した次元3

|    |  |
|----|--|
| 好  | [30, 39, 21, 86, 56, 83, 10, 87, 91, 27, 75, 22, 14, 99, 4, 20, 69, 94, 42, 88]  |
| 嫌  | [88, 30, 13, 48, 65, 69, 75, 84, 46, 20, 27, 67, 39, 77, 87, 83, 34, 91, 50, 21] |
| 褒  | [56, 39, 71, 22, 23, 21, 83, 65, 86, 10, 75, 67, 25, 30, 99, 64, 91, 92, 4, 87]  |
| 怒  | [88, 13, 46, 30, 48, 65, 84, 75, 20, 69, 87, 38, 80, 90, 77, 2, 27, 34, 0, 50]   |
| 驚P | [30, 88, 48, 13, 21, 84, 2, 18, 83, 46, 65, 39, 28, 87, 50, 27, 75, 20, 86, 91]  |
| 驚N | [88, 46, 48, 13, 30, 65, 50, 84, 77, 18, 67, 21, 80, 83, 38, 87, 90, 20, 27, 33] |
| 喜  | [39, 30, 21, 56, 91, 22, 88, 83, 27, 86, 75, 10, 69, 20, 67, 13, 71, 41, 99, 4]  |
| 悲  | [88, 48, 30, 13, 65, 46, 75, 69, 27, 39, 84, 20, 90, 21, 67, 34, 77, 80, 91, 0]  |
| 望  | [30, 56, 21, 83, 91, 71, 22, 99, 92, 10, 86, 88, 1, 4, 79, 13, 2, 39, 20, 74]    |
| 恐  | [88, 48, 13, 46, 77, 30, 84, 69, 65, 80, 38, 67, 97, 2, 90, 18, 33, 87, 75, 50]  |

表6は100次元の分散表現から, それぞれの次元からニューラルネットが予測した各感情の確率の高い順に上段右上から順に並べたものである。

One-hot ベクトルを用いたニューラルネットワークの推測を用いる手法は, 前節のニューラルネットワークの重みを用いる手法と比較した場合, 複数の感情に対して同じ次元が, 共通する要素を持つ感情に現れやすいという傾向が見られる。

とくに, [88, 30]はほとんどの感情で共通して現れており, ポジティブな感情では[21, 86, 91], ネガティブな感情では[67, 69, 77]が頻出している。しかし, これらの次元は同じ分散表現から学習させたネットワークであるにもかかわらず表4で注目されていた次元と異なっている。

表7 ネットワークが注目した次元4

|    |  |
|----|--|
| 好  | [301, 291, 228, 314, 273, 69, 308, 51, 44, 173, 120, 193, 369, 133, 344, 60, 397, 50, 151, 352]    |
| 嫌  | [53, 384, 169, 144, 323, 351, 381, 12, 77, 184, 266, 388, 24, 132, 181, 272, 155, 183, 191, 248]   |
| 褒  | [273, 193, 272, 60, 162, 223, 301, 23, 133, 270, 308, 44, 184, 195, 212, 258, 75, 243, 61, 260]    |
| 怒  | [53, 77, 169, 384, 12, 351, 381, 307, 266, 276, 184, 191, 13, 4, 45, 105, 180, 203, 274, 24]       |
| 驚P | [169, 374, 21, 384, 100, 207, 344, 144, 205, 369, 162, 164, 214, 242, 307, 340, 381, 393, 30, 388] |
| 驚N | [384, 53, 169, 381, 276, 351, 77, 97, 184, 13, 307, 394, 12, 144, 181, 374, 235, 252, 21, 171]     |
| 喜  | [193, 228, 397, 301, 258, 120, 308, 273, 314, 133, 306, 344, 20, 60, 247, 272, 291, 85, 333, 368]  |
| 悲  | [384, 12, 381, 77, 248, 307, 53, 169, 351, 132, 24, 120, 144, 184, 245, 181, 308, 272, 113, 374]   |
| 望  | [60, 86, 164, 228, 344, 301, 308, 21, 210, 127, 115, 139, 44, 223, 120, 37, 113, 273, 348, 50]     |
| 恐  | [384, 53, 203, 169, 77, 381, 12, 351, 132, 184, 307, 105, 180, 45, 144, 217, 13, 110, 274, 394]    |

同様に 400 次元の分散表現で注目されていた次元をまとめた表 7 においてもポジティブな感情では, [273, 308, 344] ネガティブな感情では, [53, 77, 169, 184, 384] が頻出するなど感情極性ごとの類似性が見られる。

また, 「好・褒」において[44, 193, 301], 「嫌・怒」において[24, 191, 266, 351, 381]が共通するなど, それらの感情は 4.3.2 項の結果と同様に関連性が高くなっていることがわかる。

#### 4.5 評価

前節までに感性情報の抽出に用いる次元を決定した。本節では得られた感性情報の評価を行う手段について説明した後実際に評価を行ってゆく。

##### 4.5.1 元の分散表現の類似語と比較

以下で示す結果は, すべて 400 次元の分散表現を用いた場合の結果である。

表 8 例 1: 「手紙」の類似語

| 比較単語  | 感情  | Word2vec | 重み    | One-hot |
|-------|-----|----------|-------|---------|
| 遺書    | 好   | 0.386    | 0.396 | 0.318   |
|       | 嫌   |          | 0.290 | 0.483   |
|       | 驚 N |          | 0.197 | 0.441   |
| ラブレター | 好   | 0.542    | 0.386 | 0.662   |
|       | 怒   |          | 0.537 | 0.498   |
| メッセージ | 好   | 0.548    | 0.665 | 0.701   |
|       | 驚 N |          | -0.01 | 0.512   |
| 記事    | 喜   | 0.316    | 0.805 | 0.265   |
|       | 望   |          | 0.402 | 0.745   |
| 書類    | 望   | 0.311    | 0.730 | 0.114   |

表 8 は「手紙」という単語と元の分散表現でのコサイン類似度が上位 100 件以内に含まれていた単語の感情ごとの類似度を求めた結果を示したものである。

たとえば, 今回例に挙げた「手紙」という単語は汎用的な語彙であり, 選択した「遺書」や「ラブレター」とは異なり感情的には中立的な単語であると考えられる。

「遺書」という単語はネガティブな感情を表現する語であると考えられ, 「嫌」や「驚 N」ではニュートラルな単語である「手紙」との類似度が低下することが予想された。ネットワークの重みを用いた手法では事前の仮説通りとなったが, ネットワークの予測結果を用いる手法では, 予想に反して類似度が上昇してしまっている。「ラブレター」の「好」に対する結果も同様である。また, 「遺書」の「好」や「ラブレター」の「怒」など一般にその感情を持たないと考えられる部分では, どちらの手法においても元の分散表現を用いた場合と近い類似度が得られている。

「手紙」の例では元の分散表現で最も類似していた単語は「メッセージ」であり, これはニュートラルに近い感情を持つ語であると考えられるが, 重みを基に感情抽出を行う手法では極端に類似度が低くなる感情が現れるなど結果に大きな差異が見られてしまった。

しかし, 全体としては「手紙」と感情の類似度が高くなる単語は「記事」や「書類」など, 同じく感情を有していないと考えられる単語が当てはまるなど類似性も見られた。

表 9 は感情を持つ単語に対して提案手法を用いた感情抽出を行った例である。この結果では「残酷」が持っている

と考えられる単語に対して元となった分散表現で類似度が高かった単語の類似度が上昇する例が数多く見られた。

表 9 例 2: 「残酷」の類似語

| 比較単語 | 感情  | Word2vec | 重み    | One-hot |
|------|-----|----------|-------|---------|
| 非情   | 嫌   | 0.436    | 0.754 | 0.152   |
|      | 怒   |          | 0.356 | 0.323   |
| 残酷   | 嫌   | 0.480    | 0.643 | 0.586   |
|      | 怒   |          | 0.733 | 0.387   |
| 残忍   | 嫌   | 0.448    | 0.625 | 0.420   |
|      | 怒   |          | 0.393 | 0.465   |
| 悲惨   | 驚 N | 0.409    | 0.619 | 0.476   |

##### 4.5.2 感情ラベルの語と比較

前項では元の分散表現の類似語と比較を行っていたが, そのある二つの単語間の距離を測るのは人間の感性に依存する。その問題を解決する手法として考えられるのがそれぞれの感情のラベルとなっている単語を尺として用いることで, 画一的な評価を行う方法である。

ただし「好」を「好き」とするなど適切に送り仮名を追加し, 「望」は「期待」と置き換えた。また「驚」は単語極性を区別せず「驚き」とした。

この時 4.1 節のアンケート結果を利用してその単語が持っている感情に対して, 「感情の類似度 > 分散表現」の類似度の場合を正答として集計を行った。その結果, 60 の単語に対して 85 の感情が想定され, いずれの精度も約 50% であった。

表 10 感情抽出の精度

|         | 正答数 | 正答率   |
|---------|-----|-------|
| 重み      | 40  | 0.470 |
| One-hot | 42  | 0.494 |

## 5. 考察

3.2.1 項のニューラルネットワークの重みを用いる手法の問題点として, ネットワークの重みや分散表現を構築する数値には負値が含まれたものが存在しており, 重みの大きな次元を単純に採用できるとは限らないという点が挙げられる。また, 感情分類モデルの学習が正しく行われていない場合, その影響を受けて感情とは無関係な次元が抽出されてしまう可能性がある。今回のネットワークにおいて正解データとなる確率が最も高くなる most 指標の精度は 0.27 程度であるため, それぞれの次元に関しても本来含まれている感情と異なる情報を含む次元が選択される可能性がある。

しかし, 4.3.2 や 4.4.1 では感情ごとに共通した次元が抽出されていたと考えられるため, 感性情報を多く含む次元が存在していることが明らかとなった。このことから, 感情分類の精度を向上させることができれば感性情報の抽出精度も向上すると考えられる。

また, 「好」「嫌」「褒」「怒」の感情に関しては, ニューラルネットワークの重みや予測に用いられていたデータにかなりの重複が見られた。今後, 用いる学習データの再検討を行うとともに, 異なる感情区分を用いて実験を行うべきと考える。

表 11 定量的評価の例

| 単語  | 感情 | Word2vec | 重み     | One-hot |
|-----|----|----------|--------|---------|
| やめる | 好  | -0.053   | 0.382  | 0.101   |
|     | 嫌  | -0.003   | -0.326 | 0.245   |
| 映画  | 好  | 0.177    | -0.018 | -0.076  |
|     | 嫌  | 0.152    | 0.125  | -0.029  |
| 完全  | 好  | -0.140   | 0.220  | -0.285  |
|     | 嫌  | -0.045   | -0.393 | -0.005  |
| 世迷言 | 褒  | 0.045    | 0.403  | 0.168   |
|     | 怒  | 0.041    | 0.061  | -0.143  |
| 二流  | 好  | -0.097   | 0.072  | 0.067   |
|     | 嫌  | -0.076   | 0.555  | 0.344   |

## 6. おわりに

本研究では Word2vec とニューラルネットワークを用いた教師有り感情分類の学習結果を利用して感情情報の抽出を行った。その結果、ニューラルネットワークが感情ごとに注目している次元に共通点が見られたものの、注目次元を用いることによる感情分類がうまく行えないことがわかった。

今後の展望としては、感性情報抽出および感情分類の精度向上を目指し、引き続き調査を行い、分類に用いる感情の区分についてさらなる検討を加えることによって、感性をより正しく表現できる分散表現の獲得を目指す。

## 謝辞

本研究は JSPS 科研費 JP20K12027, JP 21K12141 の助成を受けたものです。

## 参考文献

- [1] 青木 竜哉, 笹野 遼平, 高村 大地, “ソーシャルメディアにおける単語の一般的ではない用法の検出”, 自然言語処理, Vol.26, No.2, pp.381-406 (2019).
- [2] 中村 祐希, 乾 健太郎, “単語埋め込みの決定的縮約”, 言語処理学会第 27 次大会 (2021).
- [3] 森 祥恭, 高間 康史, “Word2Vec で表現された単語の意味の可視化に関する検討”, 2018 年度人工知能学会全国大会(第 32 回) (2018)
- [4] Peng Xu, Andrea Madotto, “Emo2Vec: Learning Generalized Emotion Representation by Multi-task Training”, Proceedings of the 9th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis, pp. 292–298, (2018).
- [5] Paul Ekman, “An Argument for Basic Emotions”, COGNITON AND EMOTION, 6(3/4), pp. 169-200, (1992).
- [6] Tomas Miklov, “Distributed Representations of Words and Phrases and their Compositionality”, Advances in Neural information Processing Systems 26, pp. 3111-3119, (2013).