

論文題目

# 文脈を考慮した一般単語の感情推定

Context-sensitive Emotion Estimation for Japanese Words

指導教授

萩原 将文 教授

慶應義塾大学 理工学部 情報工学科

令和 4 年度

学籍番号 61914694

長澤 尚武

# 目次

あらまし	1
第1章 はじめに	2
第2章 関連する研究と知見	4
2.1 感情の表現方法	4
2.1.1 Categorical Model	4
2.1.2 Dimensional Model	5
2.2 コンピュータ上での単語の表現	7
2.2.1 one-hot ベクトル	7
2.2.2 単語分散表現	8
2.3 CBOW モデルをもとにした単語感情分析手法	9
2.3.1 Continuous Bag-of-Words モデル	10
2.3.2 概要	11
2.3.3 出力例	12
2.4 BERT	14
2.4.1 特徴	14
2.4.2 BERT から得られる単語分散表現	15
第3章 文脈を考慮した一般単語感情推定手法	18
3.1 全体の概要	18
3.2 提案手法の流れ	20
3.2.1 感情の分類方法の決定	20
3.2.2 感性語の収集	20
3.2.3 感性語を含むテキストデータの取得	21

3.2.4	データセットの生成	21
3.2.5	ニューラルネットワークの学習	24
3.2.6	感情ベクトルの取得	24
<b>第4章</b>	<b>評価実験</b>	<b>25</b>
4.1	実験環境	25
4.1.1	データセット生成	25
4.1.2	ニューラルネットワークの学習条件	26
4.1.3	比較手法	26
4.2	感情ベクトルの生成例	27
4.2.1	「仕事は人生を豊かにしてくれる。」に対する出力	27
4.2.2	「仕事は面倒だし、疲れるから嫌だ。」に対する出力	28
4.3	単語の感情推定の妥当性に関する評価実験	29
4.3.1	実験設定	29
4.3.2	実験結果	29
4.3.3	考察	29
4.4	一般単語感情推定の文脈考慮性に関する評価実験	31
4.4.1	実験設定	31
4.4.2	実験結果	31
4.4.3	考察	31
4.5	感性多義語感情推定の文脈考慮性に関する評価実験	32
4.5.1	実験設定	32
4.5.2	実験結果	32
4.5.3	考察	32
<b>第5章</b>	<b>結論</b>	<b>34</b>
	謝辞	35
	参考文献	36

付録	39
付録 A 感情ベクトルの出力例	39
付録 B 実験に用いたシステム入力文	40
付録 C ニューラルネットワークの学習について	41

## あらまし

自然言語処理の分野において、対話応答システムに関する研究は盛んに行われているが、ユーザに共感し、寄り添うことのできる対話システムの実現には至っていない。それを実現するための方法として、ユーザの抱いている感情を推定するといったことがあげられる。

そこで本研究では、ユーザの発話に対してその文脈を考慮しながら、発話内に含まれる単語の感情情報を推定する手法を提案する。単語を感情に密接に関わる感性語、それ以外の一般単語にわけ、感性語との共起性をもとに一般単語の感情情報を推定する手法はすでに存在する。本手法では、ここへ文脈考慮性を導入するために、事前学習モデルの BERT から得られる分散表現を利用した。

BERT は周辺の単語の影響を加味した単語分散表現を出力することが可能である。この単語分散表現から 10 種類の感情を扱う感情ベクトルを出力するにあたり、非常にシンプルなニューラルネットワークを採用し、学習を行った。学習に用いるテキストデータそのものには感情情報に関するラベルを必要とせず、感情表現辞典により取得できる感性語とその周辺の一般単語を自動的に抽出して学習することが可能である。また、出力可能な語彙は BERT で扱うことのできる語彙と同等であるため、従来の辞書構築型アプローチに比べて対応可能語彙数は大幅に増加した。

これらの性質を利用することにより、同一単語であってもその単語が出現した文脈に応じて、取得できる感情情報を変化させることができる。評価実験により、提案手法による単語感情推定の妥当性・文脈考慮性の向上が確認された。

# 第1章

## はじめに

自然言語処理の分野において，対話応答システムの研究は盛んにおこなわれている．身の回りにあるスマートフォンや各種家電といった，様々な製品に搭載されるなど実用化も進み，人々の生活へ急速に浸透してきている．近年では，その応答性能の向上が急速に進んでおり，

また，その応答がユーザに寄り添っているようなシステムの実現に向けた研究も盛んに行われている．ユーザに寄り添った対話応答システムが実現することにより，雑談や気遣いのある応答を行うことが可能になると考えられる．人間の対話には，必ずしも特定の目的があるとは限らず，相手の興味・関心をもとに人々の生活に対話応答システムがより一層浸透し，人間とコンピュータの連携がより深く円滑に行われることで，情報化社会の進行に伴う諸問題の解決がより一層促されることが予想される．

ユーザに寄り添った対話応答システムに関する研究の方向性の一つとして，ユーザの感情推定を行うというものがあげられる．高村ら [1] の研究では，単語の感情極性，すなわち単語がポジティブとネガティブのどちらのニュアンスを持つのかを，電子のスピン方向に見立てることで推定する手法を提案している．しかし，人間が抱く複雑な感情を表現するのに2成分では情報量が少ないといえる．武内ら [2] の研究では，多数の単語に対する感情表現辞書を作成した．単語が持つ複数の感情情報を取得しているが，一つの単語に対して一対一で感情情報が保存されているため，単語がどのように使われているのかという点に対する考慮がなされていない．Kajiwara ら [3] の研究では，SNS へ投稿されたテキストに対し感情情報のラベルを付与し，投稿者の性格情報と組み合わせる

ことにより文章に対する感情情報を推定している．しかし，この研究で生成されたデータセットは 43200 件ほどであり，より高精度な文感情推定を行う上で更なる拡充を行うには，多大なコストが発生すると考えられる．

そこで本研究では，ラベルのないテキストデータからでもデータセットが生成可能な，文脈考慮性をもった単語感情推定手法を提案する．データセットの生成は，感情表現辞典 [4] を用いて感情と密接にかかわる単語（以下，感性語とよぶ）とその単語の持つ感情情報を取得し，感性語周辺の一般単語を自動的に収集することによりなされる．そのため，テキストデータに感情情報に関するラベル付けがなされていないなくても，大量の単語感情推定用データセットを生成可能である．また，事前学習モデルの BERT[5] をテキストデータの特徴量抽出機として利用した．BERT は周囲に存在する単語を考慮した文脈考慮性のある単語分散表現を出力することができる．例えば，“bank” という単語は同じスペルで“銀行”と“土手”の異なる意味を持つ．BERT では，同じ単語だが用いられている意味が異なる場合に分散表現を変化させることができる，という性質を利用した．

評価実験では，システムへ入力する文章を被験者に作成してもらったうえで，単語感情推定の妥当性，同一単語を含む文章における感情推定の文脈考慮性を評価した．

以下本論文では，第 2 章で本研究に関連する研究と知見について述べる．第 3 章では提案する文脈を考慮した一般単語感情推定手法について，第 4 章では評価実験の内容及び結果について述べる．最後に第 5 章で結論を述べる．

## 第2章

### 関連する研究と知見

#### 2.1 感情の表現方法

感情推定を行う上で，まずは感情の表現方法を定義する必要性がある．感情の表現方法の種類として，大きく2種類がある [6]．Categorical Model と Dimensional Model である．

##### 2.1.1 Categorical Model

Categorical Model とは，感情を複数のクラスで表現したものである．以下にいくつか例を示す．

###### 2.1.1.1 Ekman のモデル

Ekman[7] は怒り，嫌悪，恐れ，喜び，悲しみ，驚きの6つが普遍的な感情であり，生物学的基盤を持つと結論付けている．この基準は，異なる文化の人々の表情を読み取ることができるかどうか，である．孤立した石器時代の文化で暮らす人々が，他文化の人々の顔の表情を読み取れたかどうかという実験により確認された．

###### 2.1.1.2 Plutchik のモデル

Plutchik[8] は Ekman の6感情に加え，信頼と期待を加えた8感情のモデルを提案している．本モデルの特徴は喜びと悲しみ，信頼と嫌悪といった形で各感



情がそれぞれ対になっている点にある。

#### 2.1.1.3 中村の感情分類

中村 [4] は、感情表現辞典 [4] において、言語表現の観点から、感情を「喜、怒、哀、怖、恥、好、厭、昂、安、驚」の 10 種類に分類している。感情表現辞典では、日本の近現代作品の中から様々な感情が入り混じるような微妙な心理を描いた用例を収録している。

### 2.1.2 Dimensional Model

Dimensional Model とは、感情間の関係を表現するために次元空間を扱うようなモデルのことである。感情が連続的であるという仮説の下で表現されたものであり、一般的には 2 3 次元で表現される。主に用いられる次元は valence(感情価)、arousal(覚醒度)、dominance(優位性) の 3 つである。valence は感情が肯定的か否定的かを、arousal は興奮度を、dominance は感情に対する制御度を示している。[9][10] 以下にいくつか例を示す。

#### 2.1.2.1 Plutchik の感情の輪

Plutchik[8] は、8 つの感情をさらに 3 段階でレベル分けした、2 次元の感情モデルを提案している。次元は valence と arousal の 2 次元である。図 2.1 は、このモデルを図式化したものである。

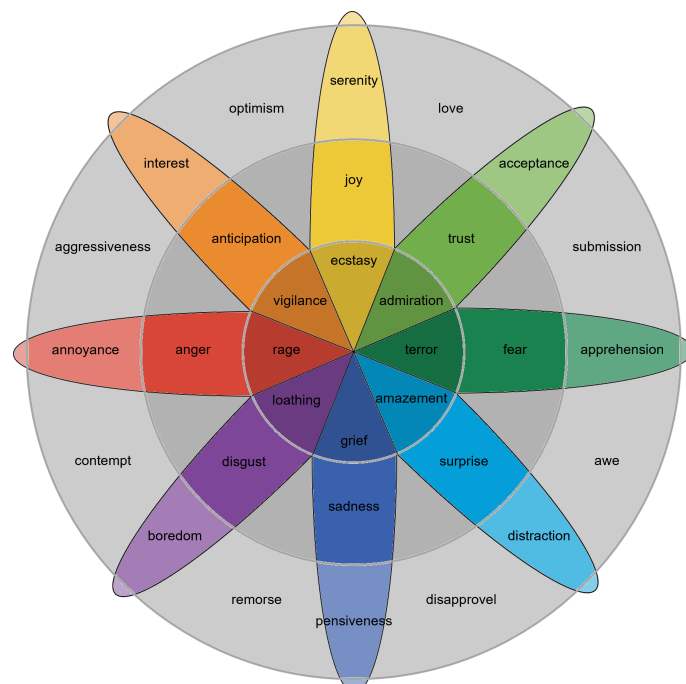


図 2.1: Plutchik の感情の輪

### 2.1.2.2 Russell のモデル

Russell[11] は感情を valence , arousal の 2 次元で表現する感情モデルを提案した . 図 2.2 は , このモデルを図式化したものである .

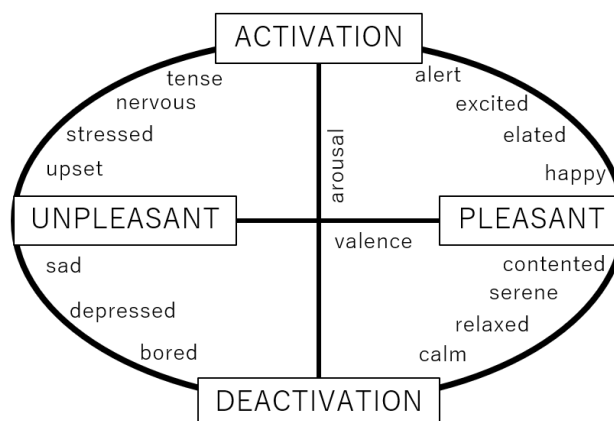


図 2.2: Russell のモデル

### 2.1.2.3 Russell, Mehrabian の PAD モデル

Rusell と Mehrabian[12] は , Pleasure , Arousal , Dominance の 3 次元で表現する感情モデルを提案した . PAD は各次元の頭文字をとったものである .

## 2.2 コンピュータ上での単語の表現

コンピュータは , 入力された単語をテキスト情報のまま処理することができない . それをかなえるためには , 単語のベクトル化が必要となる .

### 2.2.1 one-hot ベクトル

one-hot ベクトルとは , 最も単純な単語のベクトル表現方法である . 具体的には , 語彙内の単語のインデックスに対応した成分だけが 1 となり , その他の成分がすべて 0 となるようなベクトルである . よって , 次元数はシステムが扱う語彙数と同数となる . 図 2.3 は , サイズが 10000 語の語彙  $V$  について , 語彙内の単語がどのように表現されるのかを表した図である . 各単語と対応する語彙内のインデックスが括弧内の数字に対応している . 各ベクトルは , 語彙数と同じ 10000 次元である .

$$V = [a, aaron, \dots, zulu, \langle \text{UNK} \rangle] \quad |V| = 10,000$$

Man (5391)	Woman (9853)	King (4914)	Queen (7157)	Apple (456)	Orange (6257)
$\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \\ 0 \end{bmatrix}$	$\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix}$	$\begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \\ 0 \\ 0 \end{bmatrix}$	$\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix}$	$\begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$	$\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix}$
5391番目→	9853番目→	4914番目→			

図 2.3: one-hot ベクトルの概要

one-hot ベクトルによる単語ベクトルの表現を行う上での問題点として、システムの扱う語彙数が増えることにより、単語ベクトルの次元数が同量増えてしまうことが挙げられる。また、このような非常に次元の高いベクトルであるにもかかわらず、ベクトル内の1成分だけが1で他の成分はすべて0となる。よって、単語ベクトルが非常に疎なベクトルとなってしまうため、消費するリソースのわりに得られる情報量が少ないということも挙げられる。

## 2.2.2 単語分散表現

単語分散表現とは、単語をベクトル空間上の一つの点として捉えるような表現方法のことである。単語分散表現は各要素が実数値を持つ、密なベクトルとなっていて、システムの語彙数よりも低い次元数のベクトルで表現可能である。単語のベクトル同士の計算と単語間の意味関係が対応するのが特徴であり、このことからベクトルの各要素が単語の特徴を表していると考えられている。図 2.4 は、単語分散表現の特徴であるベクトル間の計算と意味関係の対応を単純化して示したものである。

	Man (5391)	Woman (9853)	King (4914)	Queen (7157)	Apple (456)	Orange (6257)
Gender	-1	1	-0.95	0.97	0.00	0.01
Royal	0.01	0.02	0.93	0.95	-0.01	0.00
Age	0.03	0.02	0.7	0.69	0.03	-0.02
Food	0.09	0.01	0.02	0.01	0.95	0.97
	$e_{Man}$	$e_{Woman}$	$e_{King}$	$e_{Queen}$		

$$e_{Man} - e_{Woman} \approx \begin{bmatrix} -2 \\ 0 \\ 0 \\ 0 \end{bmatrix}, e_{King} - e_{Queen} \approx \begin{bmatrix} -2 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

図 2.4: ベクトル同士の計算による単語間の意味関係の取得

各成分が特定の意味を持つような4次元のベクトルとして、各単語を表現している。この時、“Man”と“Woman”の違いは各ベクトルの差をとることで、“Gender”成分にあるということがわかり、同様の関係が“King”と“Queen”の間にも認められることがわかる。このように、各単語に対応するベクトルの演算結果が、単語間の意味関係と対応することになる。

なお、実際に取得される単語分散表現は、より高次元で複雑なものである。各成分が持つ意味について人間が把握することは非常に困難であるが、ベクトル間の計算と意味関係が一致する性質を踏まえると、何かしらの特徴を各成分が表現していると推測される。主要な単語分散表現の手法として、Word2Vec[13]やGloVe[14]などが挙げられる。

## 2.3 CBOW モデルをもとにした単語感情分析手法

武内ら [2] は、Continuous Bag-of-words モデルの発想を感情表現の抽出に活用した。

### 2.3.1 Continuous Bag-of-Words モデル

Continuous Bag-of-Words モデル[13](以下, CBOW モデルとよぶ)は, Word2Vec を学習するのに用いられる. Word2Vec の学習は分布仮説の下で行われる. 分布仮説とは単語の持つ意味が周囲の単語により決まる, というものである. つまり, 周囲の単語が中心の単語の意味的な要素を持っていることにもなる.

CBOW モデルは, 周囲の単語を入力して中央の単語を予測するという形のモデルである. 図 2.5 は, CBOW モデルの概要である.



図 2.5: CBOW モデルの概要

生成する分散表現の次元数を  $d$ , 語彙数を  $V$  とすると, 入力 は  $V$  次元の one-hot ベクトル, 入力と中間層の間の重みが  $(V \times d)$  の行列となる. 中間層は周囲の単語の分散表現の和となり, 出力は入力同様に  $V$  次元の one-hot ベクトルで中心の単語を出力するという形になる. この予測が正しく行えるようにネットワークを学習することで, 分散表現が単語の意味情報を反映できるようになる.

## 2.3.2 概要

### 2.3.2.1 感性語の抽出

まず，感情表現辞典 [4] から感情と密接にかかわる単語である感性語を取得する．この感情表現辞典から得られる感情情報は「喜, 怒, 哀, 怖, 恥, 好, 厭, 昂, 安, 驚」10 種類であるため，各要素が 1 つの感情に対応するような 10 次元のベクトル (以下，感情ベクトルとよぶ) を取得できる．

### 2.3.2.2 ネットワークの学習

テキストデータに対して形態素解析を行い，品詞を絞ったうえで原型に変換する．ウィンドウサイズを  $W$  としたとき，中心の感性語に対し前後  $W$  個の単語を収集し合計  $2W + 1$  個の単語群を一つのデータとする．図 2.6 で，周辺単語の収集の様子を示す．

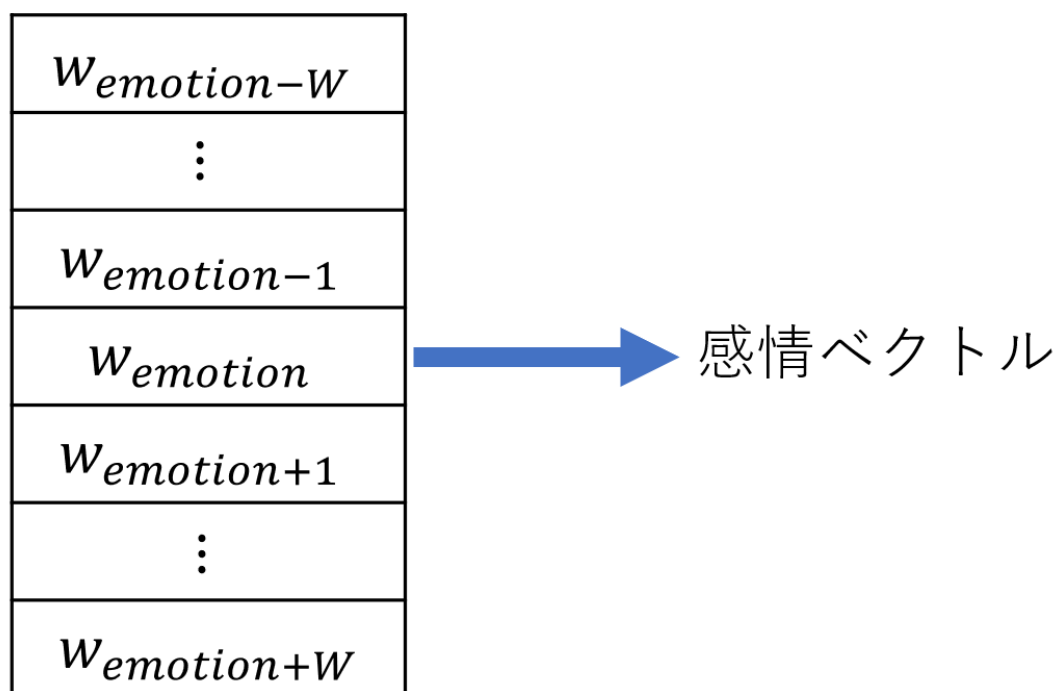


図 2.6: データセット生成での単語収集の概要

入力それぞれの単語に対する one-hot ベクトルで、中間層は感情ベクトルと同じ 10 次元のベクトルとなる。語彙数を  $V$  としたとき、入力層と中間層の間の重みは、 $(V \times 10)$  の行列で、中間層は入力単語群から変換された 10 次元のベクトルの和となる。この中間層のベクトルを softmax 層に通したときに、中心の感性語が持つ感情ベクトルとの差がなくなるようにしてネットワークを学習する。softmax 関数は以下の式で示される。

$$p_k = \frac{e^{x_k}}{\sum_{i=1}^{10} e^{x_i}} \quad (2.1)$$

softmax 関数に入力する中間層は 10 次元で、 $x_k$  はそのうちの  $k$  番目の値である。softmax 関数により、値の総和が 1 となるような同次元数の出力を得ることができる。この時、出力はウィンドウサイズ内の単語を入力した際に推定される感情の確率分布とみなされる。以上のような学習により、入力層と中間層の間の重みの各行が対応する単語に対する感情ベクトルとなる。

### 2.3.2.3 出力

上述のような学習を行うことにより、各単語に対する 10 次元の感情ベクトルを取得することができる。感情ベクトルの各成分はそれぞれ、「喜, 怒, 哀, 怖, 恥, 好, 厭, 昂, 安, 驚」の感情に対応している。これらの感情は、感情表現辞典から得られる情報に対応している。

### 2.3.3 出力例

以下に本手法で得られる出力例を示す。



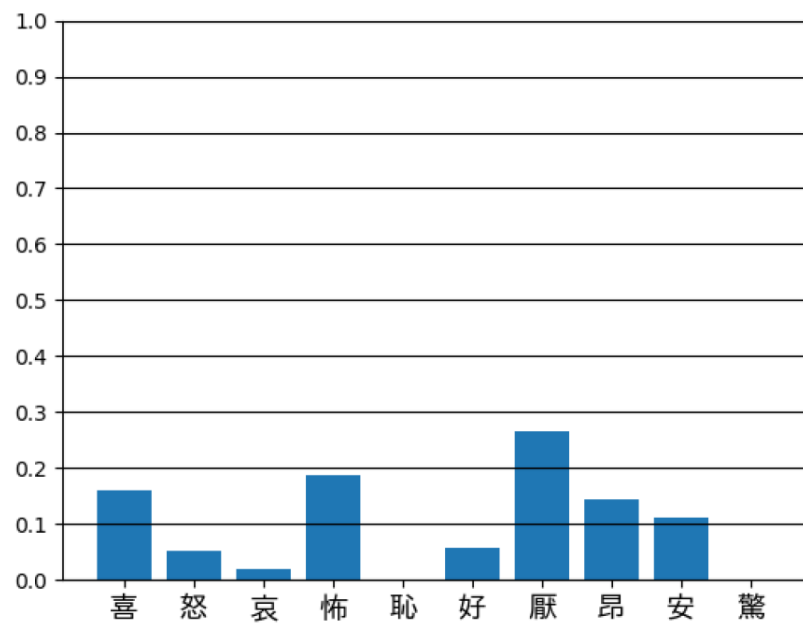


図 2.7: 本手法で得られる「投資」に対する感情ベクトルの例

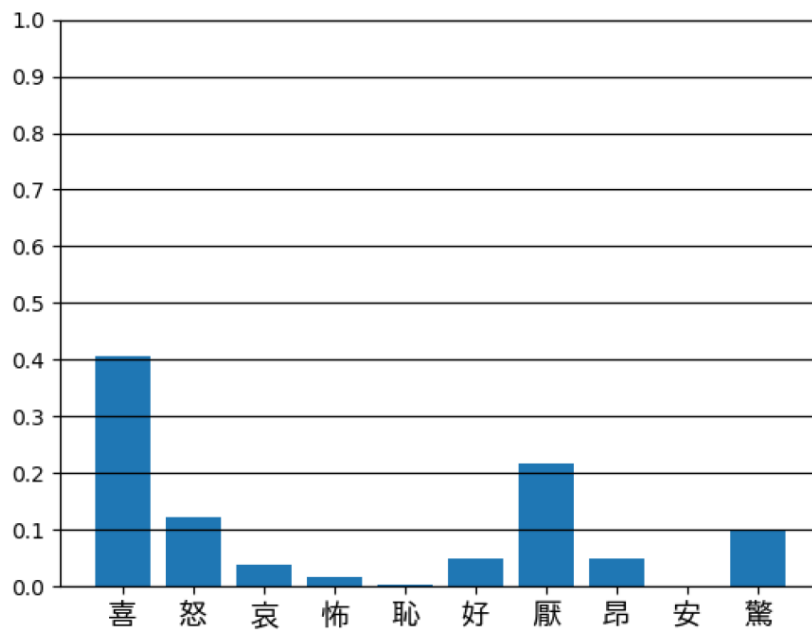


図 2.8: 本手法で得られる「成績」に対する感情ベクトルの例

本手法で得られる感情ベクトルの特徴は，一つの単語に対して，想起しうる感情が複数現れることが挙げられる．図 2.7 では「投資」という単語に対する感

感情ベクトルを示している．投資に成功して儲けが出た，という文脈においては，投資に対してポジティブな印象を持っていると考えられる．反面，投資に失敗してしまった，という文脈においては，投資に対してあまり良い印象を持っていないと考えられる．また，投資をやったことがない，という人にとっては怖いものとして扱われるケースも想定される．これらの，想定しうる「投資」のイメージを感情ベクトルがよく反映しているといえる．図 2.8 では「成績」という単語に対する感情ベクトルを示している．成績も人や試験等の結果によって抱く印象は様々であり，出力でも同様の傾向が見られている．

このように本手法では，単語が用いられる文脈によって想起されうる感情が異なってくることを感情ベクトルとして表現することができる．

本研究では，一般単語に対して想起しうる感情情報を抽出している．しかし，その情報を使って対話システム等に応用することを考えると，単語から想起しうる複数の感情が想起されるのは好ましくない．単語がどのような文脈で用いられたかを踏まえたうえでの感情推定ができれば，後続のタスクへ適切な感情情報を与えることができる．

## 2.4 BERT

Bidirectional Encoder Representations from Transformers[5](以下，BERT とよぶ) は，2018 年に Google から発表された事前学習言語モデルである．

### 2.4.1 特徴

BERT は入力として，文章をトークンに分割したものを扱い，出力として，各トークンに対応したベクトルを出力する．BERT 以前の自然言語処理で主流であった Recurrent Neural Network(以下，RNN とよぶ) ベースのモデルと同様の入出力形式である．RNN ベースのモデルでは，長い文章を入力しても最初の方の情報を保持することができないという課題があった．しかし BERT では，

Transformer Encoder[15] によりトークンの処理に他のトークンの情報を直接的に用いることができる。これに伴い、各トークンについてより文脈に即した分散表現を出力することができる。図 2.9 は、BERT の概要を図式化したものである。

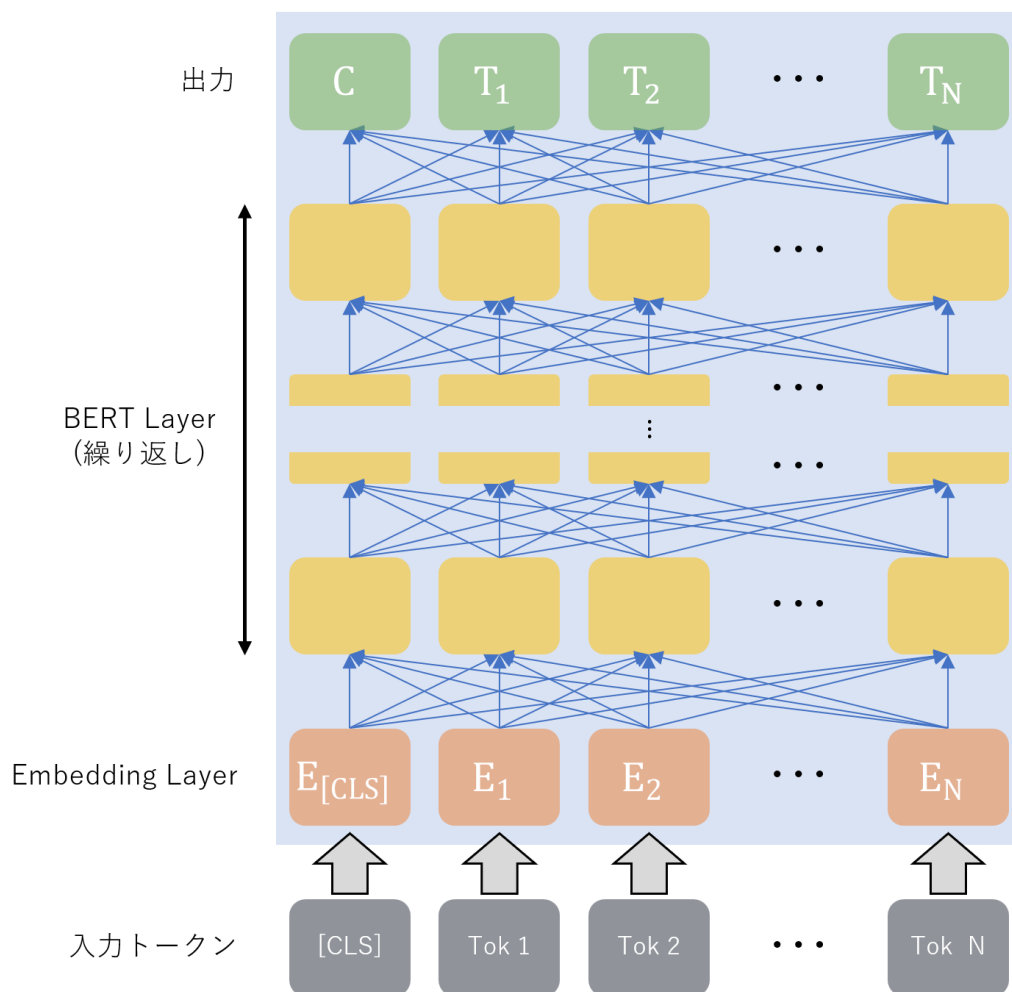


図 2.9: BERT の概要図

## 2.4.2 BERT から得られる単語分散表現

BERT から取得できる単語分散表現の特徴の一つとして、同一の単語であっても用いられる文脈によって分散表現が変化することが挙げられる。

例えば, "bank" という単語について考えたとき, まったく同じスペルでも文脈に応じて「銀行口座」という意味にも「土手」といった意味にもなる. Word2Vec 等の単語分散表現では, 単語に対して分散表現を 1 対 1 で対応付けるため, こういった多義性を表現することができず, すべて同じ出力となる. しかし BERT では, Transformer Encoder からなる Bert Layer モジュールを繰り返し通すことで単語が持つ多義性を分散表現へ反映させることができると考えられている.

以下, 分散表現が多義性を反映していることを確認する簡単な実験 [16] について述べる. "bank" という単語を用いた 3 つの文章を以下に示す.

1. I accessed the bank account.

日本語訳) 銀行口座にアクセスしました.

2. He transferred the deposit money into the bank account.

日本語訳) 彼は敷金を銀行口座に振り込みました.

3. We play soccer at the bank of the river.

日本語訳) 土手でサッカーをします.

文 1, 文 2 では "bank" を「銀行口座」という意味で用いているが, 文 3 では "bank" を「土手」という意味で用いている. 入力 Embedding Layer から得られる "bank" のトークンに一对一に対応した分散表現であるが, Bert Layer を繰り返し通すことで, この分散表現が周囲の単語の影響を受けながら変化していくことになる. 各文章の "bank" に対応する出力ベクトルの間でコサイン類似度を算出すると以下ようになる.

- 文 1 の "bank" と文 2 の "bank" の類似度: 0.8796

(どちらも「銀行口座」の意味)

- 文 1 の "bank" と文 3 の "bank" の類似度: 0.4814

(文 1 は「銀行口座」, 文 3 は「土手」の意味)

このように，“bank”という単語を異なる文章で用いたとしても，同じ意味で用いているもの同士であれば，出力ベクトルは類似した出力となる．それに対して，“bank”という単語を異なる意味で用いている文章で比較をすると，入力としては同じ単語であるにもかかわらず出力ベクトルの類似度は低下する．このような結果から，BERT から取得できる単語分散表現には文脈考慮性があるということが示唆されている．

## 第3章

# 文脈を考慮した一般単語感情推定手法

### 3.1 全体の概要

単語に対して想起される感情は，その単語が用いられる文脈に応じて変化するといえる．例えば「ドライブ」という単語について考えたとき，「海までドライブに行って気持ちよかった」という文脈で用いられた場合には，喜などのポジティブな感情を想起すると考えられる．しかし，「ドライブで事故にあってしまった」という文脈で用いられた場合には，怖などのネガティブな感情を想起すると考えられる．このように，単語に対して想起される感情は文脈によって異なる．本手法では，単語から想起される感情を文脈によって変化させる．これにより，対話システムに応用するなど，後続のタスクで使いやすい形の感情情報を与えることが期待される．

本手法では，BERT から得られる単語分散表現を活用することにより，同一単語に対しても文脈に応じて出力される感情ベクトルを変化させる．具体的には，BERT の出力する単語分散表現が文脈に応じて変化する性質を利用し，出力される感情ベクトルを変化させている．図 3.1 は提案手法の流れを示したものである．

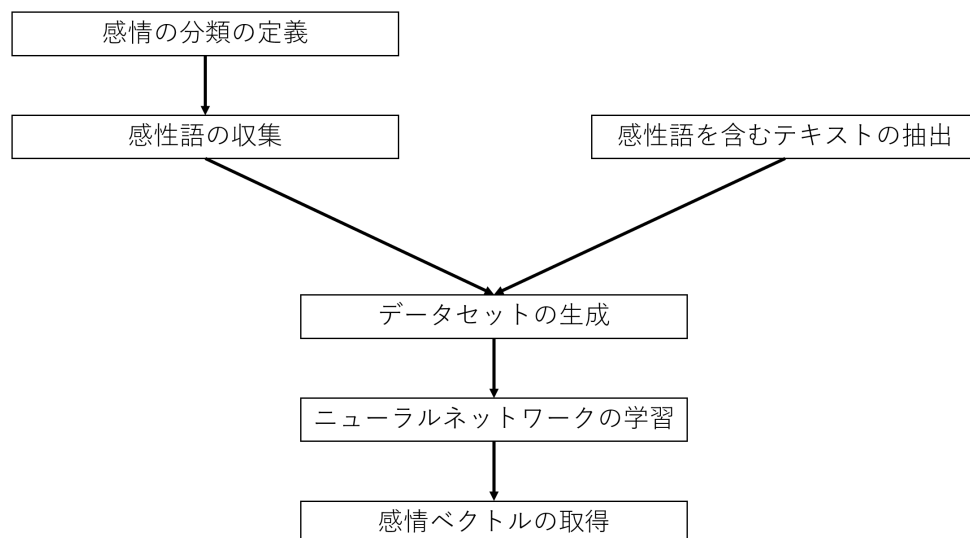


図 3.1: 提案手法の流れ

分散表現と感情ベクトルを対応付けるに当たり，感情表現辞典 [4] から取得できる単語である感性語とその感情情報を利用する．Word2Vec の学習を行う上でもとになっている分布仮説の考え方を応用し，感性語の感情が周囲の単語に影響を与えるという仮説のもとでデータセットを生成した．分散表現から感情ベクトルへの変換には非常にシンプルなニューラルネットワークを用い，学習を行った．本手法で作成したシステムでは，文章を入力することによりそれを構成している単語の感情ベクトルを入力文の文脈を考慮した上で出力することができる．また，入力が BERT の分散表現としたことで，感情ベクトルを出力可能な単語数は BERT が扱うことのできる語彙数と一致し，従来の辞書構築型アプローチと比べると対応可能語彙数は大幅に増加している．以下，本手法の流れについて述べる．

## 3.2 提案手法の流れ

### 3.2.1 感情の分類方法の決定

本手法における感情分類の方法として、感情表現辞典 [4] で取得できる感情ベクトルに合わせた 10 次元の感情分類を採用している。各感情は「喜, 怒, 哀, 怖, 恥, 好, 厭, 昂, 安, 驚」である。これらの感情は、日本語における言語表現の観点から決められたものである。

### 3.2.2 感性語の収集

感情表現辞典には、各感情を表す単語および熟語が 2278 語収録されている。各単語に対し、該当する感情に 1、該当しない感情に 0 を割り振った 10 次元の感情ベクトルを対にしてリスト化した。いくつか例を示す。

- 「めでたい」

$$(\text{喜, 怒, 哀, 怖, 恥, 好, 厭, 昂, 安, 驚}) = (1, 0, 0, 0, 0, 0, 0, 0, 0, 0)$$

- 「恐ろしい」

$$(\text{喜, 怒, 哀, 怖, 恥, 好, 厭, 昂, 安, 驚}) = (0, 0, 0, 1, 0, 0, 0, 0, 0, 0)$$

ただし、単語および熟語のうち、分かち書きで 2 単語以上に分かれてしまうものについてはリストから除外を行った。これは、分かち書きで分解された各要素のうち、どの部分に感情情報があるのかを判断するのが難しいためである。この際、分かち書きには、形態素解析エンジンの MeCab[17] を用いた。



表 3.1: 感情表現辞典から取得した感性語の例

感情	感性語の例
喜	めでたい, 歓喜, 感謝, 誇り, ...
怒	腹立たしい, 激怒, 憤り, 反感, ...
哀	悲哀, 哀悼, 嘆き, 泣き別れ, ...
怖	不気味, 恐ろしい, 震えあがる, 戦慄, ...
恥	照れる, 生き恥, 羞恥心, 赤面, ...
好	人情, 愛する, 憧れる, お気に入り, ...
厭	不快, 嫌い, 憎しみ, 残念, ...
昂	焦る, 緊迫, ときめく, 激情, ...
安	安心, 和やか, 落ち着く, 悠長, ...
驚	たまげる, 驚愕, 呆然, 以外, ...

### 3.2.3 感性語を含むテキストデータの取得

学習用データセットを作成するにあたり, 大量のテキストデータを必要とする。しかし, このテキストデータセット自体には感情にかかわるラベル付けがなされている必要はない。まず, テキストデータセットを 1 文ずつに分解したうえで, MeCab による分かち書きを行い, 文を構成する各単語の原型のリストを生成した。このリスト内に感性語が含まれているような文だけを抜粋し, 学習用データセット作成のための入力文とした。

### 3.2.4 データセットの生成

感性語を含むテキストを BERT に入力するにあたり, BERT のトークナイザで入力文をトークン化する。BERT のトークナイザは MeCab による形態素解析と wordpiece アルゴリズムによるサブワード分割により構成されている。トークン化されたテキストを BERT に入力すると, 各トークンの分散表現を取得することができる。

ここで、感性語の周辺単語の収集を行う。ウィンドウサイズを  $W$  としたとき、感性語の前後それぞれ  $W$  個の単語を収集する。なお、サブワード分割されたトークンは結合した状態でウィンドウサイズをカウントする。その際には、名詞、形容詞、形容動詞、動詞以外はカウント対象外とした。また、名詞のうち、代名詞、接尾語、非自立語、数詞についても除外を行った。除外された単語については、感情情報を持つと考えにくいという理由による。

感性語が持っている感情ベクトルをウィンドウサイズ内の単語すべてに付与することにより、学習用データセットは生成されることになる。つまり、ウィンドウサイズ内のトークンに対応する分散表現と感情ベクトルのペアが 1 件のデータということになる。以下に例を示す。

入力テキストとして、「道の真ん中で転んでしまい恥ずかしかったので、走ってその場を立ち去った」という文章を入力したとする。この文章を入力するためにトークン化を行うと、以下のように分割される。

道 / の / 真ん中 / で / 転 / ##ん / で / しまい / 恥 / ##ずか / ##しかっ / た / ので / , /  
走っ / て / その / 場 / を / 立ち / ##去っ / た / .

ここで、“##”で始まるトークンが存在するが、これは分かち書きの後さらにサブワード分割がなされたために生じるトークンである。よって、品詞による判別を行うためにサブワード分割されたトークンを結合する。すると以下のようになる。

道 / の / 真ん中 / で / 転 ##ん / で / しまい / 恥 ##ずか ##しかっ / た / ので / , /  
走っ / て / その / 場 / を / 立ち ##去っ / た / .

それぞれの分割単位で品詞を調べたうえで、ウィンドウサイズのカウント対象となるトークン列は以下の通りとなる。

道 / 真ん中 / 転 ##ん / 恥 ##ずか ##しかっ / 走っ / 場 / 立ち ##去っ

これらのトークン列のうち、「恥 ##ずか ##しかっ」の原型である「恥ずかしい」が感性語となる．よって，以下のような感情ベクトルを感情表現辞典から取得することができる．

$$(\text{喜, 怒, 哀, 怖, 恥, 好, 厭, 昂, 安, 驚}) = (0, 0, 0, 0, 1, 0, 0, 0, 0, 0)$$

ウィンドウサイズを3としたとき「恥 ##ずか ##しかっ」の前後それぞれ3つずつのトークン列へ同様の感情ベクトルを付与することになる．図3.2は，この様子を図式化したものである．

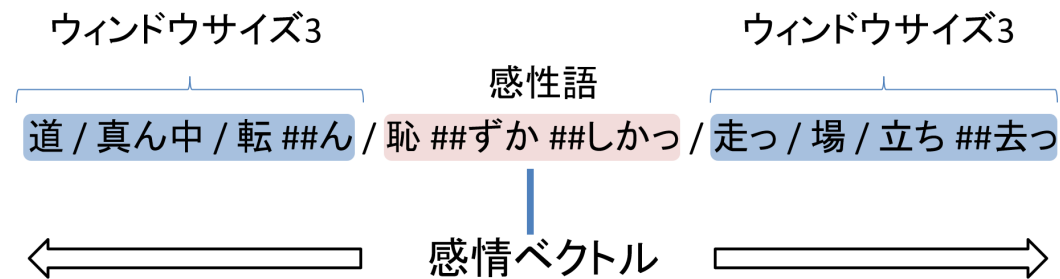


図 3.2: 中心の感性語が持つ感情ベクトルを周辺単語へ付与する様子

また本手法ではデータセット生成に当たり，与える感情ベクトルについて感性語と周辺単語の距離に応じた重み付けを行った。感性語との距離が離れるほど、与えられる感情ベクトルの強度が弱くなる。これは，感性語ではない一般単語が感性語と同等の強度でその感情を想起するわけではない，という仮説に基づくものである．表 3.2 は最小値を 0.5 とした上で感性語との距離に応じて等間隔に値が小さくなるよう感情ベクトルを付与している様子を表している．

表 3.2: 感性語からの距離に応じて感情ベクトルの強度を弱める様子

感性語との距離	-4	-3	-2	-1	0	+1	+2	+3	+4
ベクトル強度	0.5	0.625	0.75	0.875	1.0	0.875	0.75	0.625	0.5

### 3.2.5 ニューラルネットワークの学習

本手法では、BERT が出力するトークンの分散表現から感情ベクトルを出力するために、非常にシンプルなニューラルネットワークを採用し、学習を行った。図??は、用いるネットワークを図式化したものである。具体的には、入力層が分散表現、出力層となる感情ベクトルは 10 次元となるため、中間層を設けた 3 層のニューラルネットワークを学習させた。出力層には Sigmoid 関数をかけ、0 から 1 の値へ変換を施している。sigmoid 関数は以下の式で示される。

$$f(x) = \frac{1}{1 + e^{-x}} \quad (3.1)$$

### 3.2.6 感情ベクトルの取得

以上のようにして学習したネットワークに、BERT から得られたトークンの分散表現を入力することにより、入力文全体の文脈を踏まえた感情ベクトルを出力することができる。出力対象単語は、データセット作成時のウィンドウサイズカウント対象になる単語と同等の条件を満たすものに絞っている。また、サブワードに別れてしまう単語については、それぞれのトークンに対して出力される感情ベクトルの平均値をまとめて出力する。

本手法では、出力に正規化を行わなかった。その理由としては、感情表現辞典から得られる感情ベクトルには複数の感情のラベルがついた単語が存在していること、1 感情に対してだけわずかに値が出力されるようなケースではその値が極端に大きくなってしまう可能性があることが挙げられる。

## 第4章

### 評価実験

#### 4.1 実験環境

##### 4.1.1 データセット生成

感性語の抽出を行うにあたり，感情表現辞典に掲載された情報を活用する．感情表現辞典に収録されている単語と熟語の合計 2278 語のうち，MeCab で分かち書きを行うことで複数単語に別れてしまうものを除くと，1245 語となる．なお，MeCab の辞書には ipadic を用いた．本実験におけるデータセット生成では，これらを感性語とする．

データセット生成のためのテキストデータには Wikipedia[18] のテキストに前処理を施した状態で公開されている Wiki-40B[19] データセットの日本語版を用いる．1 文ずつに分割したとき，Wiki-40B データセットに含まれる文章の数は， $Train/Val/Test = 12,330,278/677,757/678,490$  となった．これらのうち，感性語を含む文は  $Train/Val/Test = 2,337,909/128,126/128,698$  となった．本実験では，ウィンドウサイズ  $W = 3$  とし，感性語の周辺単語を収集した．この BERT が出力した分散表現と感情ベクトルがペアになったデータセットのサイズは， $Train/Val/Test = 20,154,529/1,104,904/1,108,482$  となった．なお，用いた BERT モデルは Hugging Face[20] で東北大学が公開している `cl-tohoku/bert-base-japanese-whole-word-masking`[21] を用いた．

### 4.1.2 ニューラルネットワークの学習条件

入力が 768 次元の単語分散表現で，出力が 10 次元の感情ベクトルとなる．本実験では図 4.1 で示すような，中間層を 400 次元とした非常にシンプルな 3 層のニューラルネットワークを構築し，生成したデータセットを用いて学習を行った．

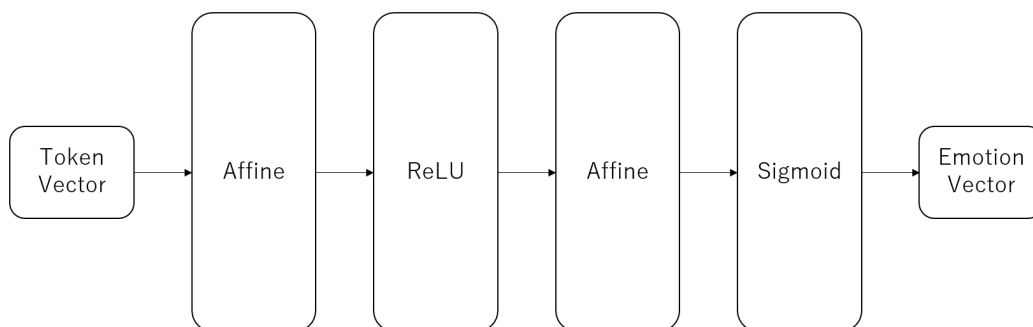


図 4.1: ニューラルネットワークの構造

中間層の活性化関数は ReLU 関数とし，出力層に sigmoid 層を通すことにより，0 以上 1 以下の値に変換する．この時，出力された 10 次元のベクトルが，データセットに与えられた感情ベクトルに近づくよう学習を行うことになる．損失関数は最小二乗誤差とし，Adam により学習率  $1 \times 10^{-3}$  で最適化を行った．エポック数については，Early Stopping で 5 回以上 validation loss が更新されなかった段階で学習を終了させた．

### 4.1.3 比較手法

本手法と比較するためのベースライン手法として，BERT Layer を通す前の Embedding Layer による分散表現をネットワークの入力とした場合を設定する．BERT Layer を通すことにより，Transformer による周りの単語の影響を加味することになる．今回設定したベースラインでは Embedding Layer による分散表現を利用することで周辺単語の影響が含まれないことになる．また，感性語の周辺単語に与える感情ベクトルについて距離に応じた重み付けを行うことによ

る効果を確認するため、重み付けを行う BERT+weight、行わない BERT に分けて実験を行う。

## 4.2 感情ベクトルの生成例

入力文に対する各単語の感情ベクトル出力例を以下に示す。

### 4.2.1 「仕事は人生を豊かにしてくれる。」に対する出力

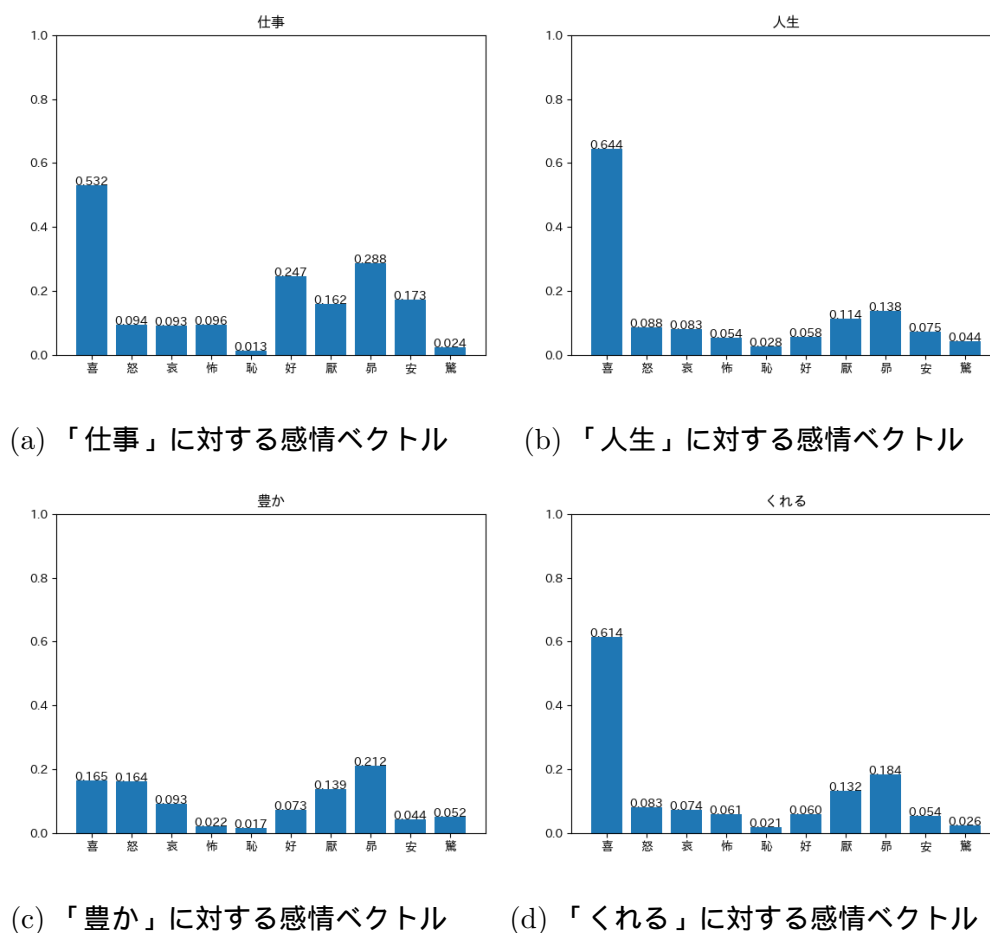


図 4.2: 「仕事は人生を豊かにしてくれる。」に対する各単語の感情ベクトル

## 4.2.2 「仕事は面倒だし、疲れるから嫌だ。」に対する出力

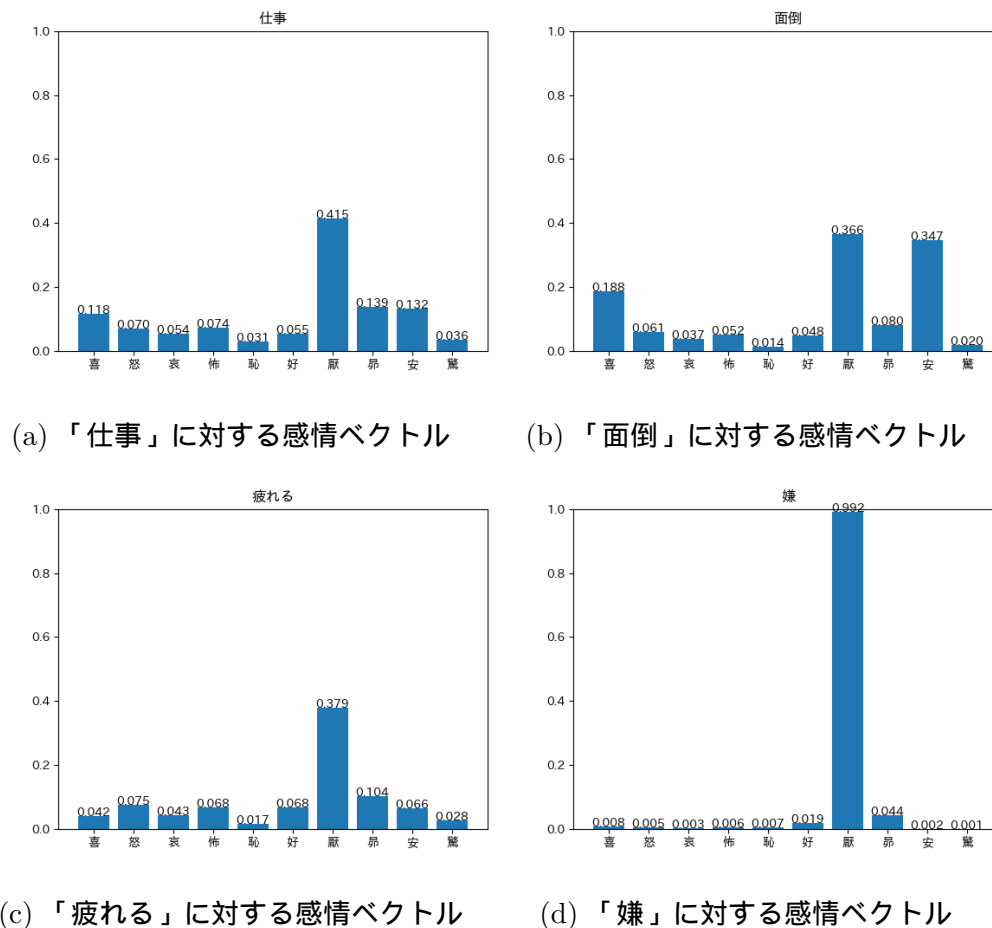


図 4.3: 「仕事は面倒だし、疲れるから嫌だ。」に対する各単語の感情ベクトル

図 4.2 ではポジティブな文脈で「仕事」という単語を用いているのに対し，図 4.3 ではネガティブな文脈で「仕事」という単語を用いている．それぞれの「仕事」に対する感情ベクトルを比較すると，文脈に応じてその出力が変化していることがわかる．



## 4.3 単語の感情推定の妥当性に関する評価実験

### 4.3.1 実験設定

本実験では，本システムにより出力された単語の感情ベクトルが人間の想起する感情のイメージと合致しているのかを検証する．実験に先立ち，被験者には特定の単語を用いるような文章を生成してもらい，システムに入力したうえで，文章生成時に指定した単語（以下，対象単語とよぶ）の感情ベクトルを出力した．対象単語について，それが用いられた文脈を考慮した上で想起されるような感情を感情ベクトルの形で被験者に予想してもらった．システム出力，被験者により予想された感情ベクトルをそれぞれ正規化したうえで，Top-1 精度，Top-3 精度，コサイン類似度を比較する．同次元のベクトル  $\mathbf{a}, \mathbf{b}$  について，これらのコサイン類似度は以下のように計算される．

$$\cos(\mathbf{a}, \mathbf{b}) = \frac{\mathbf{a} \cdot \mathbf{b}}{\|\mathbf{a}\| \|\mathbf{b}\|} \quad (4.1)$$

### 4.3.2 実験結果

表 4.1: 実験 1 の結果

	Top-1	Top-3	cos 類似度
BERT+weight	32.95%	61.36%	<b>0.612</b>
BERT	<b>37.50%</b>	<b>65.91%</b>	0.609
embedding	26.14%	53.41%	0.557

### 4.3.3 考察

これから分析を行います．

- Top-k 精度では，感性語との距離に応じて付与するベクトル強度に重み付けを行った BERT+weight よりも，重み付けを行っていない BERT のほうが高性能．

- cos 類似度では BERT+weight が最高性能

これらの結果から，最も強く想起する感情を判定すると，強度の高いベクトルを与えた BERT のほうが，強度を弱めている BERT+weight よりも高性能が出る．しかし，ベクトル全体の距離を cos 類似度で測ると，BERT+weight が高性能であることから，複数の感情が想起されるケースなどでは BERT+weight のほうがより良い出力を行っている可能性がある．

このような流れでまとめていこうと考えています．

表 4.2: 単語の種類別で分けた実験 1 の Top-1 精度

	一般単語	感性語	感性多義語
BERT+weight	<b>27.78%</b>	<b>61.11%</b>	23.53%
BERT	25.00%	<b>61.11%</b>	<b>38.24%</b>
embedding	11.11%	50.00%	29.41%

表 4.3: 単語の種類別で分けた実験 1 の Top-3 精度

	一般単語	感性語	感性多義語
BERT+weight	<b>58.33%</b>	<b>77.78%</b>	55.88%
BERT	<b>58.33%</b>	<b>77.78%</b>	<b>67.65%</b>
embedding	33.33%	72.22%	64.71%

表 4.4: 単語の種類別で分けた実験 1 のコサイン類似度

	一般単語	感性語	感性多義語
BERT+weight	<b>0.608</b>	<b>0.728</b>	0.555
BERT	0.595	0.715	<b>0.567</b>
embedding	0.521	0.663	0.538

単語の種類別で精度を出した時の特徴についてもまとめていこうと考えています．

## 4.4 一般単語感情推定の文脈考慮性に関する評価実験

### 4.4.1 実験設定

本実験では、同一の単語に対する出力がその出現する文脈に応じて適切に変化するのかを検証する。特定の単語（以下、対象単語と呼ぶ）を含み、文脈がポジティブ・ネガティブになるような文章をそれぞれ被験者に生成してもらい、システムに入力した。本手法により得られた対象単語の感情ベクトルについて、それぞれの出力で、適切な感情を出力できているか、不適切な感情の出力を抑制できているかを、また対象単語について、用いられる文脈が異なっている2つの出力を比較して、文脈を考慮した出力になっているかをそれぞれ1(最低) 5(最高)の5段階評価で評価してもらった。

### 4.4.2 実験結果

表 4.5: 実験2の結果

	適切な感情の出力	不適切な感情の出力の抑制	文脈考慮性
BERT+weight	3.93	<b>3.82</b>	<b>3.92</b>
BERT	<b>3.98</b>	3.76	3.91
embedding	3.22	2.88	2.60

### 4.4.3 考察

これから分析を行います。

- 適切な感情の出力で BERT が最高性能，不適切な感情の抑制と文脈考慮性で BERT+weight が高性能

このことから，周辺単語について距離を弱める BERT+weight のほうが，不適切な感情を抑えることができている可能性があると考えています。

## 4.5 感性多義語感情推定の文脈考慮性に関する評価実験

### 4.5.1 実験設定

本実験では、データセット生成に用いた感情表現辞典において、多くの感情を想起するとされている、多義性のある感性語（以下、感性多義語とよぶ）に対しても、文脈を考慮して適切な感情を出力することができるのかを検証する。本実験では、4 つ以上の感情を想起するとされていた感性多義語の中から、「気持ち」、「涙」、「思い」の3 単語について、文脈考慮性を持った単語感情推定ができるのかを検証する。特定の感性多義語（以下、対象単語とよぶ）を含み、文脈が異なっている文章をあらかじめ被験者に生成してもらい、システムに入力した。本手法により得られた対象単語の感情ベクトルについて、それぞれの出力で適切な感情を出力できているか、不適切な感情の出力を抑制できているか、をそれぞれ1(最低) 5(最高) の5 段階評価で評価してもらった。

### 4.5.2 実験結果

表 4.6: 実験 3 の結果

	適切な感情の出力	不適切な感情の出力の抑制
BERT+weight	4.05	<b>3.84</b>
BERT	<b>4.12</b>	3.48
embedding	3.66	3.31

### 4.5.3 考察

これから分析を行います。

- 適切な感情の出力では BERT が、不適切な出力の抑制では BERT+weight が高性能と、一般単語同様の傾向が見られる。

このことから，一般単語の場合と同様の形でまとめる予定です．

## 第5章

### 結論

実験結果の分析を踏まえて記述します。

## 謝辞

本研究を行うにあたり親身に相談に乗っていただき，ご指導してくださった萩原将文教授，ならびに共に問題解決，議論，相談に付き合ってくださった研究室の先輩方，同期の皆様に深く感謝いたします。誠にありがとうございました。

## 参考文献

- [1] 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”. In: 情報処理学会論文誌 47.2 (Feb. 2006), pp. 627–637. ISSN: 1882-7764. URL: <https://cir.nii.ac.jp/crid/1050001337882589440>.
- [2] 武内 達哉, 萩原 将文. “単語の持つ感情推定法の提案と単語感情辞書の構築”. In: 日本感性工学会論文誌 18.4 (2019), pp. 273–278. DOI: 10.5057/jjske.TJSKE-D-18-00104.
- [3] Tomoyuki Kajiware et al. “WRIME: A New Dataset for Emotional Intensity Estimation with Subjective and Objective Annotations”. In: *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*. Online: Association for Computational Linguistics, June 2021, pp. 2095–2104. DOI: 10.18653/v1/2021.naacl-main.169. URL: <https://aclanthology.org/2021.naacl-main.169>.
- [4] 中村 明. 感情表現辞典. 東京堂出版, 1993.
- [5] Jacob Devlin et al. *BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding*. 2018. DOI: 10.48550/ARXIV.1810.04805. URL: <https://arxiv.org/abs/1810.04805>.
- [6] Hande Aka Uymaz and Senem Kumova Metin. “Vector based sentiment and emotion analysis from text: A survey”. In: *Engineering Applications of Artificial Intelligence* 113 (2022), p. 104922. ISSN: 0952-1976. DOI: <https://doi.org/10.1016/j.engappai.2022.104922>. URL: <https://www.sciencedirect.com/science/article/pii/S0952197622001452>.



- [7] Paul Ekman. “An argument for basic emotions”. In: *Cognition and Emotion* 6.3-4 (1992), pp. 169–200. DOI: 10.1080/02699939208411068. eprint: <https://doi.org/10.1080/02699939208411068>. URL: <https://doi.org/10.1080/02699939208411068>.
- [8] Robert Plutchik. “A psychoevolutionary theory of emotions”. In: *Social Science Information* 21.4-5 (1982), pp. 529–553. DOI: 10.1177/053901882021004003. eprint: <https://doi.org/10.1177/053901882021004003>. URL: <https://doi.org/10.1177/053901882021004003>.
- [9] Rafael Calvo and Sunghwan Kim. “Emotions in text: Dimensional and categorical models”. In: *Computational Intelligence* early view (Sept. 2012). DOI: 10.1111/j.1467-8640.2012.00456.x.
- [10] SREEJA P S and Mahalakshmi G S. “Emotion Models: A Review”. In: *International Journal of Control Theory and Applications* 10 (Jan. 2017), pp. 651–657.
- [11] James Russell. “A Circumplex Model of Affect”. In: *Journal of Personality and Social Psychology* 39 (Dec. 1980), pp. 1161–1178. DOI: 10.1037/h0077714.
- [12] James A Russell and Albert Mehrabian. “Evidence for a three-factor theory of emotions”. In: *Journal of Research in Personality* 11.3 (1977), pp. 273–294. ISSN: 0092-6566. DOI: [https://doi.org/10.1016/0092-6566\(77\)90037-X](https://doi.org/10.1016/0092-6566(77)90037-X). URL: <https://www.sciencedirect.com/science/article/pii/009265667790037X>.
- [13] Tomas Mikolov et al. *Efficient Estimation of Word Representations in Vector Space*. 2013. DOI: 10.48550/ARXIV.1301.3781. URL: <https://arxiv.org/abs/1301.3781>.
- [14] Jeffrey Pennington, Richard Socher, and Christopher D. Manning. “GloVe: Global Vectors for Word Representation”. In: *Empirical Methods in Nat-*

- ural Language Processing (EMNLP)*. 2014, pp. 1532–1543. URL: <http://www.aclweb.org/anthology/D14-1162>.
- [15] Ashish Vaswani et al. *Attention Is All You Need*. 2017. DOI: 10.48550/ARXIV.1706.03762. URL: <https://arxiv.org/abs/1706.03762>.
- [16] 小川 雄太郎. *つくりながら学ぶ! PyTorchによる発展ディープラーニング*. マイナビ出版, 2019.
- [17] Taku Kudo, Kaoru Yamamoto, and Yuji Matsumoto. “Applying Conditional Random Fields to Japanese Morphological Analysis”. In: *Proceedings of the 2004 Conference on Empirical Methods in Natural Language Processing*. Barcelona, Spain: Association for Computational Linguistics, July 2004, pp. 230–237. URL: <https://aclanthology.org/W04-3230>.
- [18] *Wikipedia 日本語版*. <https://ja.wikipedia.org/wiki/メインページ>.
- [19] Mandy Guo et al. “Wiki-40B: Multilingual Language Model Dataset”. In: *LREC 2020*. 2020. URL: <http://www.lrec-conf.org/proceedings/lrec2020/pdf/2020.lrec-1.296.pdf>.
- [20] *Hugging Face*. <https://huggingface.co/>.
- [21] *cl-tohoku/bert-base-japanese-whole-word-masking*. <https://huggingface.co/cl-tohoku/bert-base-japanese-whole-word-masking>.

## 付録 A

### 感情ベクトルの出力例

本手法で得られる感情ベクトルの画像をいくつか掲載する予定です。

## 付録 B

### 実験に用いたシステム入力文

評価実験を行うにあたり，研究室の皆様にご生成いただいた入力用文章の一覧を掲載する予定です．

## 付録C

### ニューラルネットワークの学習について

ウィンドウサイズや活性化関数を変更したときの性能変化と出力の特性についてまとめる予定です。また、学習の進行状況 (Loss が落ちる様子) も掲載予定です。