

BERTによる支援記録文章の感情極性推定の一手法

山本 隆太[†] 松本 典久^{††} 上野 史^{†††} 太田 学^{†††}

[†] 岡山大学工学部情報系学科 〒700-8530 岡山市北区津島中 3-1-1

^{††} 岡山大学大学院自然科学研究科 〒700-8530 岡山市北区津島中 3-1-1

^{†††} 岡山大学学術研究院自然科学学域 〒700-8530 岡山市北区津島中 3-1-1

E-mail: [†]{p0pm8dxf, piia6ope}@s.okayama-u.ac.jp, ^{††}{uwano, ohta}@okayama-u.ac.jp

あらまし 高齢者や障がい者が過ごす福祉支援施設では、入所者が起こすトラブルであるインシデントの発生は問題なため、インシデントの予兆をとらえて発生を未然に防ぐことが重要である。本研究では、インシデントの予兆が入所者の支援記録に現れると考え、その記録の本文が内包する感情の極性を推定する。具体的には Bidirectional Encoder Representations from Transformers(BERT) をファインチューニングしたものをを用いて支援記録文章の感情極性を分類し、実験によりその推定精度を評価する。感情極性推定実験では、支援記録文章に Positive, Negative, Both, Others の 4 種類からなる感情極性ラベルを付与し作成した支援記録文章のコーパスと、ホテルレビュー文に Positive, Negative, Neutral の 3 種類の感情極性ラベルが付与されている、株式会社リクルートの公開している JRTE コーパスを用いる。実験の結果、支援記録文章の感情極性推定は、ファインチューニングに用いる支援記録文章の件数を 200 件から 800 件まで増加させた際に精度の上昇が大きかった。また最大の 2,000 件の支援記録文章でファインチューニングした場合が最も精度がよく、精度の向上の余地があると分かった。支援記録文章と JRTE コーパスの実験の比較により、文章内に含まれる特有の表現が推定精度に影響を及ぼしている可能性があることが分かった。

キーワード 感情極性推定, BERT, 医療・ヘルスケア

1 はじめに

近年、高齢者や障がい者が入所する福祉支援施設の需要が高まってきている。福祉支援施設では、入所者の人数に対する職員の数は限られており、支援には限度がある。そのため、入所者による器物破損や他者への傷害などといったインシデントが施設内で発生すると、業務の妨げとなり、入所者は十分な支援を受けられなくなる。しかし、インシデントを事前に察知することはベテランの職員でも難しく、入所者が起こすインシデントの予兆を発見し未然に防止することは重要な課題である。

本研究では、インシデントの予兆が入所者の記録に現れると考え、その記録本文の内包する感情極性を推定する¹。感情極性についての研究は人との対話に関わる分野で多く行われており、その有用性が示されている。本研究では福祉支援施設という人との関りが重要な場所での問題を取り扱うため、感情極性推定が予兆検出に資すると考えた。具体的には、福祉支援施設の支援記録のうち、記録本文にあたる文章の感情極性を推定し、推定精度を評価する。推定には、自然言語処理モデルの 1 つである Bidirectional Encoder Representations from Transformers(BERT)[1] を利用する。また感情極性推定実験では、支援記録の文章のみではなく、JRTE コーパスも用いて実験を行い、実験結果を比較する。

本稿の構成を述べる。2 節では関連研究を紹介する。3 節で

は、本実験で使用する支援記録文章と感情極性推定器について説明する。4 節では、感情極性推定実験の内容と結果を示し、考察を述べる。5 節では、まとめと今後の課題について述べる。

2 関連研究

2.1 BERT

Bidirectional Encoder Representations from Transformers (BERT)[1] は、Devlin らが提案した自然言語処理モデルであり、双方向 Transformer[2] というニューラルネットワークを利用している。BERT は大量の教師なしデータによる事前学習を行ったモデルを、比較的少量の教師ありデータでファインチューニング (FT) して個々のタスクに合わせて学習する。これにより事前学習済みモデルをそのまま適用するよりも性能の向上が期待できる。

BERT の特徴として、文章を文頭と文末の双方向から学習することで、文脈を考慮できることが挙げられる。BERT ではこの双方向からの学習を、Masked Language Model (MLM) と Next Sentence Prediction (NSP) の 2 つのタスクを行うことで可能としている。MLM では、まず入力文章を単語ごとにトークンに分割し、その中からランダムに選ばれた 15 % のトークンを確率的に別のトークンで置き換える。具体的には、選択された 15 % のうち、80 % を [MASK] という特殊なトークンに、10 % をランダムな別のトークンに置き換え、残りの 10 % はそのままのトークンとする。その後 [MASK] トークンに置き換えられたトークンを予測し学習するのが MLM

1: 本研究は、岡山大学の研究倫理審査専門委員会の承認を得て実施した。(研 2104-005)

のタスクである。NSP は 2 つの入力文を受け取り、そのペアにおいて 2 つ目の文が 1 つ目の文の後続の文となっているかを予測するタスクである。具体的には、文の片方を 50 % の確率で他の文に置き換え、それらが隣り合っているか隣り合っていないか判別することで学習する。BERT は公開時に、General Language Understanding Evaluation (GLUE) ベンチマーク [3] の 8 つのタスク、Stanford Question Answering Dataset (SQuAD)[4] の v1.1 と v2.0, Situations With Adversarial Generations (SWAG)[5] の計 11 の自然言語処理タスクで最高精度を記録した。

2.2 感情極性推定

徳久らは、ユーザの感情を推定可能な対話システムの実現のために、感情推定手法を提案した [6]。徳久らは、まず約 130 万件からなる感情生起要因コーパスを構築し、そのコーパスを利用して 2 段階の感情推定を行った。第 1 段階として、入力されたユーザ発話の感情極性を Positive と Negative に分類し、第 2 段階として、第 1 段階で分類した感情極性値と大量の感情生起要因を利用し、ユーザ発話の感情を推定する。評価実験の結果、2 段階の感情推定手法は 1 段階の感情推定手法に対して有用であることが分かり、特に対話システムで致命的な誤応答の原因となる感情極性の誤りが有意に低減されることが分かった。

池川らは、Twitter のツイートなどに含まれる顔文字の持つ役割に着目し、ツイートの感情極性推定を行った [7]。池川らは、まず言語と顔文字の感情極性の組み合わせがツイートの感情極性に与える効果を分析し、顔文字の役割を提案した。その後、顔文字の役割を利用したツイートの感情極性分析を行った結果、顔文字の役割を利用した場合の方が、多くのツイートの感情極性を正しく推定することができた。

中澤らは、BERT を用いた単文の感情極性推定手法を提案した [8]。中澤らは、読売新聞のニュース記事から抽出した各文がポジティブであるかネガティブであるかを示す感情極性値を定義し、提案手法である BERT を用いた感情極性の推定手法、構文解析や形態素解析を用いた感情極性推定法、及びアンケートから得られた感情極性値の比較を行った。その結果、BERT を用いた提案手法は、教師データの中にポジティブなデータが少なかったため、アンケートから得られた感情極性値よりもネガティブと誤推定される場合が多くなり、今後の教師データの拡充や調整などの工夫が必要であることが分かった。

[6], [7], [8] のような感情極性の研究は、インシデントの予兆検出にも役立つと考え、本研究では、感情極性推定をする。

2.3 インシデント検出

松本らは、BERT を用いた支援記録の分類モデルを提案し、その分類性能を実験により評価した [9]。この提案モデルは支援記録の本文や日付情報などを入力として受け取り、入力された記録をインシデントか非インシデントかの 2 クラス、またはインシデントの種類に応じた「暴力のみ」「器物破損のみ」「暴力かつ器物破損」「インシデントなし」の 4 クラスに分類している。実験結果として、松本らの提案手法は SVM, k-近傍法、

ランダムフォレストのいずれの手法も上回る分類精度を記録した。また、彼らは入力データとして記録本文に別の情報を加えることで精度の向上が見込めることを示唆している。

本研究では、インシデントの分類を行っている [9] の研究とは異なり、インシデントそのものではなく、記録本文が内包する感情極性を推定する。また本研究により推定された感情極性はインシデントの予兆検出に利用できる可能性がある。

3 支援記録と感情極性推定

3.1 支援記録

感情極性推定に用いる支援記録は、福祉支援施設の職員が各入所者への支援について記録したものであり、入所者、入力者、処理日時、記録本文からなる。このうち本研究で用いるのは、記録本文に記述されている文章であり、入所者の様子や活動内容等が記述されている。この文章は、文の長さや文字数、記述形式といったものが定まっていない自由記述形式の文章となっている。本稿では、この支援記録に記述されている記録本文のことを支援記録文章と呼ぶ。

3.2 支援記録文章と感情極性

まず感情極性を定義し、支援記録文章にラベルを付与した。ラベルの付与は本稿の第一著者が手で行い、感情極性ラベルは Positive, Negative, Both, Others の 4 種類とする。具体的には、施設入所者の機嫌が良い、落ち着いている様子が記録されている文章は Positive、施設入所者の機嫌が悪い、情緒が乱れている様子が記録されている文章、または他者への暴力行為などが記録されている文章は Negative、記録内に前述の Positive と Negative の両方の特徴が含まれる文章は Both、上記 3 つのどれにも当てはまらない文章は Others としている。ラベル付けされた支援記録文章の具体例を表 1 に示す。(a) の例では、文章内に「元気」や「ご機嫌」といった入所者の機嫌が良いと取れる表現があるため、Positive としている。(b) の例では、「機嫌が悪い」という表現が文章内に含まれているため Negative としている。(c) の例では、「穏やか」、「ご機嫌」といった Positive の特徴を持つ表現と「機嫌が悪い」や「暴言」といった Negative の特徴を持つ表現の両方が文章中に存在するため Both としている。(d) の例では、午前と午後に行った活動が記録されており、Positive と Negative のどちらの要素にもあたらないため、Others としている。

3.3 感情極性分類器

感情極性分類器は BERT の事前学習済みモデルをファインチューニングすることで作成する。本稿では、東北大学乾研究室が公開している訓練済み日本語 BERT モデルの Version2²を使用する。その事前学習済みモデルに対して、支援記録文章にラベル付けを行ったコーパスを用いてファインチューニングすることにより、感情極性分類器を作成する。この感情極性分類器は文章を入力することで、その文章の感情極性が分類結果と

2: <https://www.nlp.ecei.tohoku.ac.jp/news-release/3284/>

表 1 支援記録文章とその感情極性の例

	支援記録文章	感情極性
(a)	終日元気に参加される。ご機嫌でよく職員に話しかけていた。	Positive
(b)	福祉園にて帽子が無くなったと大騒ぎをした様子。帰ってからも若干機嫌が悪い。	Negative
(c)	穏やかに過ごされていたが、他者にぬいぐるみを盗られてから機嫌が悪くなってしまい、暴言が多くなってしまった。しかし、入浴はご機嫌だった。	Both
(d)	AM: リング通しを少しされその後は外で過ごされていた。 PM: 内科検診のため帰園される	Others

して出力される。

4 感情極性推定実験

4.1 実験概要

4.1.1 実験に用いるコーパス

感情極性推定実験には福祉支援施設の入所者 1 名についての支援記録文章 2,416 件を用いる。この支援記録文章には 3.2 節で説明した 4 種類の感情極性ラベルが付与されており、その感情極性ラベル付き支援記録文章で感情極性分類器をファインチューニングして感情極性推定実験を行う。また、このデータセットの内、無作為に選んだ 2,000 件をファインチューニング (FT) に用いる FT 用データ、残りの 416 件を評価に用いるテスト用データとする。図 1 に支援記録文章の感情極性分布を示す。

また実験では、Japanese Realistic Textual Entailment Corpus (JRTE コーパス) を使った感情極性推定実験も行う。JRTE コーパスは株式会社リクルートが公開しているコーパスであり、じゃらん net に投稿されたホテルレビュー文 5,553 件で構成されている。各レビュー文は開発用、学習用、テスト用に分けられており、Positive, Negative, Neutral の 3 種類の感情極性ラベルが付与されている。実験では開発用、学習用データから無作為に選んだ 2,000 件を BERT の FT 用データとし、テスト用データは元の 553 件をそのまま用いる。図 2 に使用する JRTE コーパスの感情極性分布を示す。

4.1.2 実験内容

本研究では、以下の (1)-(3) の実験を行う。

- (1) FT に用いる支援記録文章の件数を変えた感情極性推定実験
- (2) 支援記録文章の 4 種類の感情極性のデータ件数を揃えた感情極性推定実験
- (3) 支援記録文章と JRTE コーパスを用いた感情極性推定精度の比較

まず実験 (1) では、BERT の FT に用いる支援記録文章のデータ件数を 200~2,000 の間で 200 件ずつ変化させ、感情極性を推定しその精度を評価する。この実験に用いる FT 用支援記録文章の内訳を表 2 にまとめる。

実験 (2) では、支援記録文章の FT 用データとテスト用データにおいて 4 種類の感情極性のデータ件数を揃えた場合とそうでない場合を比較する。具体的には、表 3 に示すように作成した支援記録文章のうち、最も件数の少ない感情極性ラベルに合わせて、他のラベルのデータ件数を減らすことで、データ件数を揃える。実験 (2) では、FT 用とテスト用の両方のデータ件

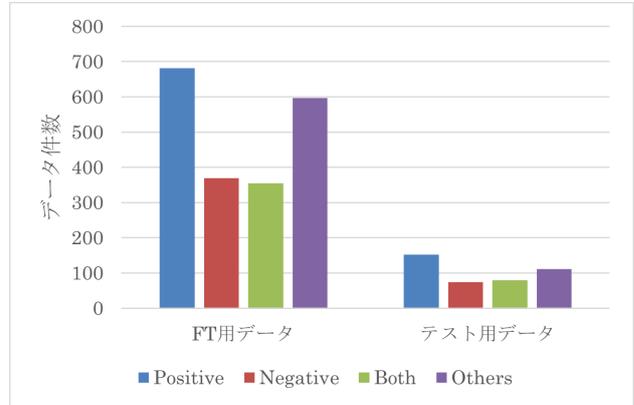


図 1 支援記録文章の感情極性分布

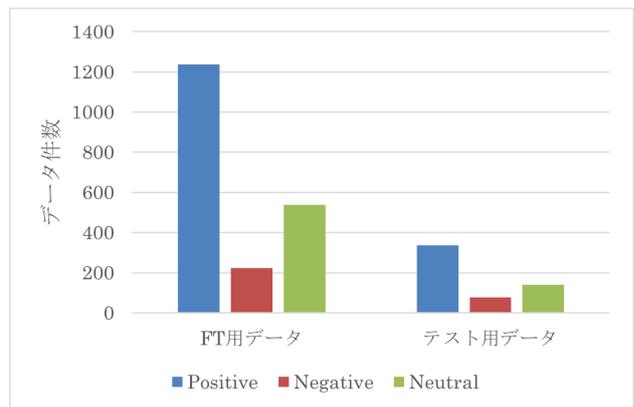


図 2 JRTE コーパスの感情極性分布

表 2 実験 (1) で使用する FT 用支援記録文章の感情極性ラベルの内訳

各感情極性ラベルのデータ件数 (割合)				
Positive	Negative	Both	Others	合計
63(31.50%)	44(22.00%)	33(16.50%)	60(30.00%)	200
127(31.75%)	86(21.50%)	75(18.75%)	112(28.00%)	400
208(34.67%)	116(19.33%)	109(18.17%)	167(27.83%)	600
280(35.00%)	142(17.75%)	144(18.00%)	234(29.25%)	800
356(35.60%)	181(18.10%)	172(17.20%)	291(29.10%)	1,000
424(35.33%)	211(17.58%)	209(17.42%)	356(29.67%)	1,200
488(34.86%)	246(17.57%)	251(17.93%)	415(29.64%)	1,400
548(34.25%)	290(18.13%)	286(17.88%)	476(29.75%)	1,600
618(34.33%)	324(18.00%)	320(17.78%)	538(29.89%)	1,800
681(34.05%)	369(18.45%)	354(17.70%)	596(29.80%)	2,000

数を揃えた場合、FT 用データのみ件数を揃えた場合、データの件数を揃えない場合の 3 つの場合について実験する。ただしデータの件数を揃えない場合は、両方のデータ件数を揃えた場合と合計のデータ件数が等しくなるように、元の支援記録文章

表 3 実験 (2) で使用する支援記録文章の感情極性ラベルの内訳

	種別	データ件数 (割合)				合計
		Positive	Negative	Both	Others	
両データを揃える	FT	350(25.00%)	350(25.00%)	350(25.00%)	350(25.00%)	1,400
	テスト	70(25.00%)	70(25.00%)	70(25.00%)	70(25.00%)	280
FT 用データのみ揃える	FT	350(25.00%)	350(25.00%)	350(25.00%)	350(25.00%)	1,400
	テスト	101(36.07%)	52(18.57%)	55(19.64%)	72(25.71%)	280
どちらのデータも揃えない	FT	488(34.86%)	246(17.57%)	251(17.93%)	415(29.64%)	1,400
	テスト	101(36.07%)	52(18.57%)	55(19.64%)	72(25.71%)	280

表 4 実験 (3) で使用するコーパスの感情極性ラベルの内訳

コーパス	種別	各感情極性ラベルのデータ件数 (割合)			
		Positive	Negative	Neutral	合計
支援記録文章	FT	681(34.05%)	369(18.45%)	950(47.50%)	2,000
	テスト	152(36.54%)	74(17.79%)	190(45.67%)	416
JRTE コーパス	FT	1,238(61.90%)	223(11.15%)	539(26.95%)	2,000
	テスト	336(60.76%)	78(14.10%)	139(25.14%)	553

のコーパスから無作為にデータを減少させる。

実験 (3) では、JRTE コーパスと支援記録文章の両方を用いて感情極性を推定し、その結果を比較する。支援記録文章は、感情極性ラベルの種類を JRTE コーパスに揃えるため、Both と Others を統合し Neutral とする。FT、テスト用共に支援記録文章を用いた場合、FT、テスト用共に JRTE コーパスを用いた場合、FT 用に JRTE コーパス、テスト用に支援記録文章を用いた場合の 3 つの場合について実験を行い、推定精度を評価する。実験 (3) で用いるコーパスの内訳を表 4 にまとめる。

また全ての実験について FT の際のパラメータは epoch 数 10、バッチサイズ 8 である。

4.1.3 評価指標

感情極性推定実験の評価には、再現率 (Recall)、適合率 (Precision)、F 値 (F-measure) を利用する。まず Positive, Negative, Both, Others の 4 つの各感情極性ラベルについてこれらの値を求める式は以下に示す。

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (1)$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (2)$$

$$\text{F-measure} = \frac{2\text{Recall} * \text{Precision}}{\text{Recall} + \text{Precision}} \quad (3)$$

例えば、感情極性が Positive 場合の TP, TN, FP, FN は下記のようなになる。

- TP: Positive だと分類された文章のうち、実際に Positive である文章の件数
- FP: Positive だと分類された文章のうち、Positive ではない文章の件数
- TN: Positive ではないと分類された文章のうち、実際に Positive でない文章の件数
- FN: Positive ではないと分類された文章のうち、Positive である文章の件数

よって再現率は全ての Positive である文章に対し、分類器が正しく Positive と分類したものの割合であり、適合率は分類器が Positive と分類した文章のうち、実際に Positive であるものの割合である。また F 値は再現率と適合率の調和平均となる。

評価では 4 つの各感情極性ラベルについて個別にこの値を計算した後、その平均を求める。また実験 (3) では、Both, Others ではなく Neutral についてこれらの評価指標を算出する。

4.2 実験結果

4.2.1 FT に用いる支援記録文章の件数と感情極性推定精度

FT に用いる支援記録文章のデータ件数を変えて感情極性推定を行った結果の再現率、適合率、F 値を図 3 に示す。この図は横軸が FT に用いた支援記録文章のデータ件数、縦軸が各指標の値を示しており、グラフの値は 4 種類の感情極性ラベルの平均の値である。この結果より、FT に用いるデータ件数が多くなるにつれ、どの指標も値が上昇していることが分かる。特にデータ件数が 200 件から 800 件の間では、上昇幅が大きい。またデータ件数が 800 から 2,000 件の間においても精度の上昇が見られるものの、200 件から 800 件の間と比べるとその上昇は緩やかとなっている。しかし、再現率、適合率、F 値の全ての指標について最も良い値を記録しているのは FT 用データ件数が 2,000 件の場合であるため、まだデータの追加により精度向上が期待できる。

4.2.2 支援記録文章の 4 種類の感情極性ラベルのデータ件数を揃えた場合の感情極性推定精度

支援記録文章の 4 種類の感情極性ラベルのデータ件数を揃えた場合の感情極性の推定結果の F 値を図 4 に示す。4 種類の感情極性ラベルとその平均について、表 3 に示した 3 つの場合の F 値を示している。結果を見ると、Negative 以外のラベルではデータの件数を揃えない場合に最も F 値が高くなっている。Negative についても両データの件数を揃えた場合とデータ件数を揃えない場合は、ほとんど変わらない F 値となっている。平均についてもデータの件数を揃えない場合がわずかではあるが最も F 値が高くなっている。

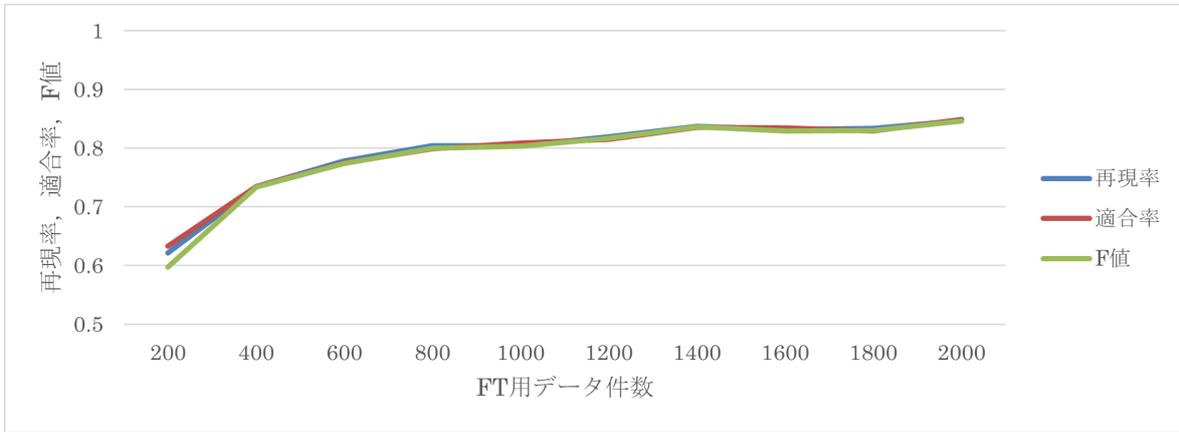


図3 FT用支援記録文章の件数と4ラベルの平均感情極性推定精度

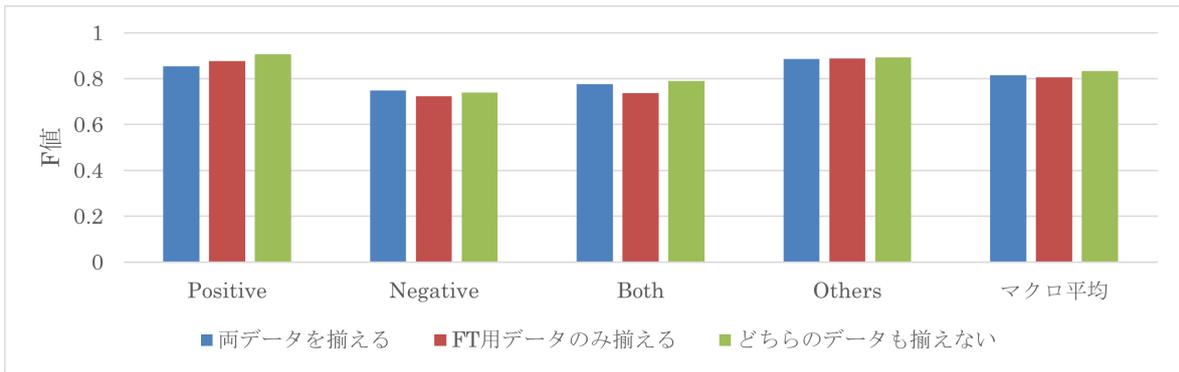


図4 4種類の感情極性ラベルのデータ件数を揃えた場合の感情極性推定精度

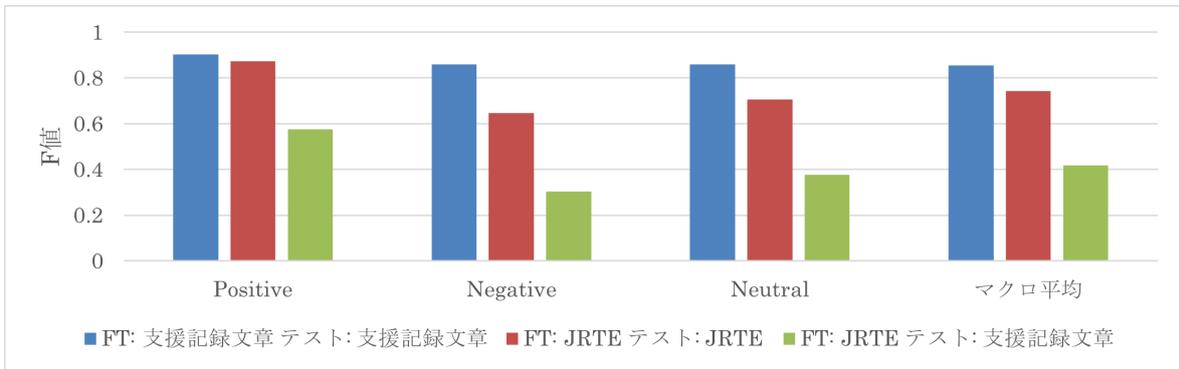


図5 支援記録文章とJRTEコーパスを用いた場合の感情極性推定精度

表3を見ると、データ件数を揃えなかった場合とFT用データのみ件数を揃えた場合でPositiveは488件から350件、Othersは415件から350件と調整を行った場合の方が件数が少なくなっている。そのためPositiveとOthersについてはFT用データ件数を揃えた場合に推定精度が下がるという予想通りの結果になっている。NegativeとBothのFT用データについては、Negativeが246件から350件、Bothが251件から350件とデータを揃えた場合の方が件数が多い。しかし推定精度はほとんど変化しない、もしくは低下するという結果となった。これよりNegativeとBothの結果について、必ずしも学習データにおいてそのラベルのデータ件数が少ないことだけが推定精度が低いことの原因ではないことが分かる。

4.2.3 支援記録文章とJRTEコーパスを用いた場合の感情極性推定精度

支援記録文章とJRTEコーパスの2つのコーパスを用いて感情極性を推定した。その結果のF値を図5に示す。Positive, Negative, Neutralの3つの感情極性ラベルとその平均について、表4で示した3つの場合のF値を示している。結果を見ると、全てのラベルと平均において、FT、テスト用共に支援記録文章を用いた場合が最も高い値となっている。またFT用にJRTEコーパス、テスト用に支援記録文章を用いた場合のF値が最も低くなった。

表4に示す内訳を見ると、Positiveについては、JRTEコーパスのデータ件数の方が支援記録に比べて2倍近くも多いにも

表 5 JRTE コーパスの文章と感情極性の例

文章	感情極性
次回もまた利用したいと思います。	Positive
言うことなしですね。	Positive
駅直結で便利なホテルです。	Positive
安いから仕方ないのかな?	Negative
本当に残念です。	Negative
態度が悪い。	Negative
迷うことはありません。	Neutral
下呂温泉は初めてでした。	Neutral
個人的な意見ですが。	Neutral

かわらず, JRTE コーパスの Positive の推定精度の方が低い結果となっている. JRTE コーパス含まれる文章例を表 5 に示す. 支援記録文章と違い, JRTE コーパスは単文で構成されている. そのため文脈を考慮できるという BERT の特徴を十分に生かしきれない上, 各文に含まれる情報が少ない. また JRTE コーパスは様々なホテルのレビュー文から構成されており, 1 人の人物について数名の入力者が記述をした支援記録文章と比べて, 文章に一貫性がない. これらの理由から支援記録文章を用いた場合と比べて FT が難しく, JRTE コーパスを使用した場合の方が推定精度が低くなった可能性がある.

4.3 誤推定事例の分析

実験 (1) で FT 用データを 2,000 件用いた場合において, 誤推定となった事例を表 6 に示す. まず (a) の文章について, 文章内に「情緒が不安定」との記述があるため, 正解ラベルは Negative であるが, 推定結果は Others となった. この文章は記録文章が比較的長く, そのほとんどを感情の伴わない活動の記録, つまり Others の特徴を持つ文章が占めている. そのため, 最初に「情緒が不安定」という Negative な記述が出てきたものの, Others に誤推定された可能性がある. 次に (b) の文章は, 「ニコニコ」という Positive に該当する記述と「暴言」, 「大声」といった Negative に該当する記述の両方が存在するため, この正解ラベルは Both であるが, 推定結果は Negative となった. (b) の文章には, Negative の要素は「情緒悪く」, 「暴言」, 「大声」と多いのに対して, Positive の要素が「ニコニコ」という部分しかないため, これが原因の可能性もある. この「ニコニコ」という表現は「ニコニコと笑顔で」というように「笑顔」や「笑う」といった表現と共に使用されていることが多く, 「ニコニコ」のみでは文脈によって Positive として十分に認識されない可能性がある.

これらの事例のうち (b) の事例について, 「ニコニコ」という語が Positive として認識されているか, また「笑顔」という類似語と比べ認識に差があるのかを確認する実験を行った. 具体的には, 表 6 の (b) の文章の「ニコニコされながら」という部分を「笑顔を見せながら」という表現に変更した (b)' の感情極性を推定する. 加えて「ニコニコしている。」という「ニコニコ」という表現のみ含む文である (c) を用意し, 感情極性を推定した. なお (b)' の文章の正解の感情極性ラベルは Both, (c) の文章の正解の感情極性ラベルは Positive である. 推定結果

を表 7 に示す. 表 7 より「ニコニコされながら」という部分を「笑顔を見せながら」に変更した場合, 推定結果が正しく Both となることが分かる. また「ニコニコしている。」という文も正しく Positive と推定されている. このことより, 「ニコニコ」という表現はそれのみでも Positive と認識されるよう学習がされているが, 文章の文脈によっては, 正しく推定ができない場合があると分かる. 一方「笑顔」という表現は「ニコニコ」という表現よりも, Positive と認識されやすいといえる. この差について使用した支援記録文章 2,416 件を全て確認すると, 「ニコニコ」は 34 件の文章に使用されているのに対して, 「笑顔」は 191 件の文章に使用されていた.

5 まとめ

本研究では, 福祉支援施設の支援記録文章が内包する感情極性の推定手法を提案し, その推定精度を実験により評価した. 提案手法は, 日本語版 Wikipedia で事前学習された BERT を支援記録文章でファインチューニングして用いるもので, 入力された文章をその感情極性に依拠して Positive, Negative, Both, Others の 4 ラベルに推定する.

実験では, 福祉支援施設の入所者 1 名についての支援記録文章 2,416 件にこの 4 種類の感情極性ラベルを付与した. またホテルレビュー文に Positive, Negative, Neutral の感情極性ラベルが付与されている, 株式会社リクルートの公開している JRTE コーパスも実験に用いた. 実験では, ファインチューニングに用いる支援記録文章の件数を変えて感情極性推定精度を評価する実験 (1), 支援記録文章の各感情極性ラベルのデータ件数を揃えて感情極性推定精度を評価する実験 (2), 支援記録文章と JRTE コーパスを用いて感情極性推定精度の比較を行う実験 (3) の 3 つの実験を行った. 実験 (1) では, ファインチューニングに用いるデータ件数が多いほど精度向上につながるが, 実験で使用したデータの上限である 2,000 件近くでは件数の増加による精度上昇は小さいことが分かった. 実験 (2) では, 作成したコーパスに存在するラベルのデータ件数の少なさが推定精度の低下の原因ではないことが示唆された. 実験 (3) では, 推定結果やコーパスに含まれる文章を比較したところ, 文章の記述の一貫性や文章特有の表現が推定精度に影響を及ぼす可能性があることが分かった.

今後は, 支援記録文章に感情極性ラベルを付与して新たなコーパスを作成するとともに, 推定精度の低かったラベルの文章について分析を進め, 推定精度の向上に繋がる手法を模索したい. また文章記述の一貫性や特有の表現が推定精度に及ぼす影響についても分析を進める予定である.

文 献

- [1] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova, “Bert: Pretraining of deep bidirectional transformers for language understanding,” arXiv preprint arXiv:1810.04805, 2018.
- [2] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Lukasz Kaiser, and Illia Polosukhin, “Attention is all you need,” Advances in Neural

表 6 感情極性推定を誤った事例

	支援記録文章	正解の感情極性	推定結果
(a)	夕食時、最初は情緒が不安定である為か「ご飯はいらん。」と言う。何度か尋ねるも同じ反応をされる。数分経った後ご飯も何も持たずに声をかけると「お薬飲む。」と言われる。すぐにお茶を提供し、お薬を飲んでいただくことができる。その後、本人の居室へご飯を持って行くと「ご飯食べる。」と言われる。最終的には、おかずを少し残してご飯を食べることができる。	Negative	Others
(b)	帰園後から情緒悪く、暴言が多い。入浴前も大声があったが、お風呂にはすんなり来てくださり入浴中はニコニコされながら入られていた。お風呂から上がった後は再び暴言・大声がある。	Both	Negative

表 7 「ニコニコ」という表現を含む文の感情極性推定の結果

	支援記録文章	正解の感情極性	推定結果
(b)'	帰園後から情緒悪く、暴言が多い。入浴前も大声があったが、お風呂にはすんなり来てくださり入浴中は笑顔を見せながら入られていた。お風呂から上がった後は再び暴言・大声がある。	Both	Both
(c)	ニコニコしている。	Positive	Positive

Information Processing Systems 30, pp. 5998–6008, 2017.

- [3] Alex Wang, Amanpreet Singh, Julian Michael, Felix Hill, Omer Levy, and Samuel R. Bowman, “GLUE: A Multi-Task Benchmark and Analysis Platform for Natural Language Understanding,” arXiv preprint arXiv:1804.07461, 2018.
- [4] Pranav Rajpurkar, Jian Zhang, Konstantin Lopyrev, and Percy Liang, “SQuAD: 100,000+ Questions for Machine Comprehension of Text,” arXiv preprint arXiv:1606.05250, 2016.
- [5] Rowan Zellers, Yonatan Bisk, Roy Schwartz, and Yejin Choi, “SWAG: A Large-Scale Adversarial Dataset for Grounded Commonsense Inference,” arXiv preprint arXiv:1808.05326, 2018.
- [6] 徳久良子, 乾健太郎, 松本裕治, “Web から獲得した感情生起要因コーパスに基づく感情推定,” 情報処理学会論文誌, Vol.50, No.4, pp.1365–1374, 2009.
- [7] 池川知里, 新妻弘崇, 太田学, “顔文字の役割を利用したツイートの感情極性推定,” DEIM2014, E6-42, 2014.
- [8] 中澤政孝, 亀井且有, 前田陽一郎, クーパーエリック, “BERT を用いた単文の感情極性推定手法の提案とその有効性,” FSS2020, TA2-3, 2020.
- [9] 松本典久, 上野史, 太田学, “福祉支援施設の支援記録を利用したインシデント発生検出の一手法,” DEIM2022, A34-1, 2022.