

## 進捗報告

### 1 今週行ったこと

- 研究発表会資料のために、多値分類タスクの結果等の取りまとめをした。
- 研究発表会の資料作成をした。

### 2 はじめに

昨今の AI の発展は目まぐるしく、様々な応用分野でその有用性が示されている。しかしその一方で、機械学習手法はその判断根拠が不明な点が問題視されているため、近年では人工知能に判断根拠の説明を示させる「説明可能な人工知能」に関する研究が盛んである。自然言語処理の分野においても先述の研究は盛んであるため、このような研究に取り組むことを目標とし、本研究ではラベルが付与された日本語レビュー文章の 2 値分類や多値分類を通して BERT モデルの理解や周辺知識の習得に努めた。

### 3 要素技術

#### 3.1 BERT

BERT-[1] は 2018 年 10 月に Google の Devlin らの論文で発表された自然言語処理モデルである。Transformer というアーキテクチャを組み込み、文章を双方向から学習することによって文脈を読み取ることが実現された事前学習モデルである。ファインチューニングをすることで様々な自然言語処理タスクに対応することが出来る汎用性の高さが注目を集めた。本研究の実験では、東北大学の乾・鈴木研究室によって公開されている BERT 日本語 Pretrained モデルを使用した。

### 4 関連研究

#### 4.1 アスペクトベースの感情分析

アスペクトベースの感情分析-[2] では文章に含まれるアスペクト情報を利用することで、その文章がどのような事柄について書かれたものかを分析することが出来る。アスペクト情報-[3] とはその文章の

カテゴリを表しており、文章が何を対象としているかを示すエンティティと、その対象のどの属性について言及しているかを示すアトリビュートによって定義される。アスペクトベースの感情分析では 3 つのステップで分析を行う。1 つ目は与えられたアスペクトカテゴリに文章进行分类する。2 つ目では文中に含まれるアスペクトカテゴリに対するフレーズの位置を推定する。3 つ目でフレーズの極性を分析し、その精度を向上させることを目指す。

### 5 データセット

#### 5.1 評判分析用チェックデータ

実験 1 では日本語レビュー文章とそれぞれの文章のラベルが与えられたデータを用いた。ラベルは、レビュー文章がネガティブなら 0、レビュー文章がポジティブなら 1 として与えられている。総データ数は 6000 であった。

#### 5.2 楽天トラベルレビュー：アスペクトセンチメントタグ付きコーパス

実験 2 では日本語レビュー文章とそれぞれの文章のラベルが 12 個与えられたデータを用いた。総データ数は 76624 で、そのうち全ラベルが 0 であるデータは 28255 であり、今回はこれらを除くことで少なくとも 1 つのラベルが立っているデータのみを用いた。その結果として総データ数は 48369 となった。

### 6 実験

本研究では 2 値分類で日本語文章データのポジネガタスクを解き、精度評価を行った。実験 1 では評判分析用チェックデータを用いた 2 値分類を行った。BERT モデルの末尾にネガポジ分類のための全結合層を追加し、出力として 2 クラス分類 [ ネガティブ ( 0 ) or ポジティブ ( 1 ) ] を出力するモデルを用いた。クラス分類には入力した文章データの 1 単語目 [ CLS ] の特徴量を利用した。また、BERTLayer の最終層と全結合層のみ fine-tuning を行った。

## 6.1 実験1

実験1では評判分析用チェックデータを用いた2値分類をした。BERTの最大入力長は512トークンなのでそれを超える文章データは取り除いた。表1に該当するデータを取り除いた後の実験に用いたデータの内訳を示す。表2にそれぞれのデータに含まれるポジティブラベル数とネガティブラベル数を示す。表3に実験時のパラメータについて示す。

表1: 2値分類に用いたデータの内訳

総データ数	訓練データ数	テストデータ数
5638	4511	1127

表2: ポジティブ, またはネガティブのラベルが付与されているデータの数

	ポジティブ数	ネガティブ数
訓練データ	2350	2161
テストデータ	573	554

表3: 2値分類タスクに用いたパラメータ

値	パラメータ
入力層の次元数	768
出力層の次元数	2
バッチサイズ	32
最適化関数	Adam
損失関数	CrossEntropyLoss

クラス分類には文章データの1単語目[CLS]の特徴量を識別器の入力として用いた。上記の訓練データを用いて5分割検証を行い, 5個のモデルを作成した。表4にそれぞれのモデルごとの正解率を示す。また, 表5には5個の中で最も正解率の高かったモデルを用いて表1のテストデータでの正答率, 再現率, F1値, を求めた結果を示す。そして, 表6には5個の中で最も正解率の高かったモデルとテストデータを用いた実験におけるコンフュージョンマトリクスを示す。そして, 5分割検証を行って正解率の高いモデルを得ることがどれほど有効なのかを調べるために交差検証を行わず, 表1の訓練データをそのまま学習させ, テストデータを用いて実験を行った。表7に交差検証を行わず, 訓練データをそ

のまま学習した場合の正答率, 再現率, F1値を示す。表8には交差検証を行わず, 訓練データをそのまま学習した場合のコンフュージョンマトリクスを示す。最後に, 表9には交差検証時にテストデータを用いた実験における予測を間違えたデータの具体例を示す。図1と図2を比較すると, 交差検証を行って正解率の高かったモデルを用いる方がより正しくデータを分類できていることがわかる。

表4: 2値分類タスクにおける5分割交差検証での正解率

0.8208	0.8554	0.8642	0.8483	0.8669
--------	--------	--------	--------	--------

表5: 最も正解率の高かったモデルを用いたテストデータでの正解率, 再現率, F1値

正解率	再現率	F1値
0.8669	0.8865	0.8713

表6: 交差検証を行わない場合のテストデータでの正解率, 再現率, F1値

正解率	再現率	F1値
0.8279	0.9058	0.8425

表7: 5分割検証をした場合でのコンフュージョンマトリクス

		予測の分類結果	
		ポジティブ	ネガティブ
実際の分類結果	ポジティブ	469	85
	ネガティブ	65	508

表8: 訓練データをそのまま用いた場合でのコンフュージョンマトリクス

		予測の分類結果	
		ポジティブ	ネガティブ
実際の分類結果	ポジティブ	414	140
	ネガティブ	54	519

## 6.2 実験2

実験2では楽天トラベルレビュー: アスペクトセンチメントタグ付きコーパスを用いた多値分類をし

表 9: 予測を間違えたデータの具体例

テキスト	ラベル	予測ラベル
<p>朝夕に富士山が見えてこそ、の料金だと思いました。見えなければ、高すぎます。</p> <p>宿からの回答: この度は当館をご利用頂き誠に有難うございました。富士山をご覧になれず大変残念でございました。</p> <p>富士山をご覧になれなくともお客様にご満足頂ける旅館を目指し</p> <p>努力して参りたいと思います。どうぞまたのお越しをお待ち申し上げます。</p>	0	1
<p>金沢駅裏で立地は最高です。ホテル内もきれいでフロントも好感がもてます。</p> <p>部屋の設備は普通レベルで価格相応だと思います。これまでに経験したことがなかったのですが、</p> <p>ベッドの堅さと枕の具合が悪くたびたび目を覚まし、2泊とも熟睡できませんでした。</p>	0	1
<p>結婚式で利用しましたが、スタッフの方が優しく、安心して結婚式を任せる事ができました。</p> <p>式を予約してから式までホテルのフィットネスクラブが無料で使えるので、</p> <p>結婚式を予定している人は早めに予約して長く利用するとお得かも !?</p>	1	0
<p>食事等では何度も利用していましたので、使い勝手の良い立地はよく知っていました。</p> <p>宿泊はすごく高いだろうと思っていたら意外と手頃なプランがあるのですね。</p> <p>新宿西口方面のシティホテルは駅から遠いものが多いなか、ここは駅から徒歩でも苦にならない距離なのでいいですね。</p> <p>地下街もありますし。さすがに古い感じはしますが、サービスも悪くないし、チャチャラしていない老舗ホテルの貫禄があります。</p>	1	0

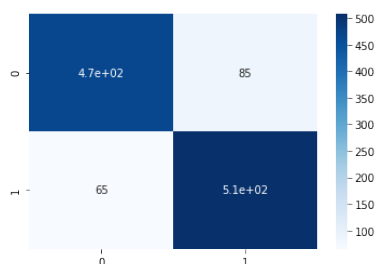


図 1: 5 分割検証をした場合でのコンフュージョンマトリクス

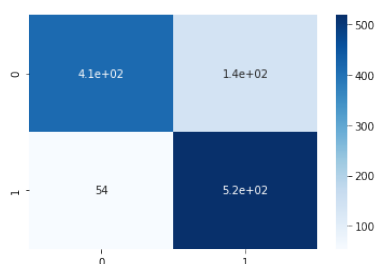


図 2: 訓練データをそのまま用いた場合でのコンフュージョンマトリクス

た. 表 10 に実験で用いたデータの内訳を示す. また, 表 11 に実験で用いたパラメータを示す. クラス分類には文章データの 1 単語目 [CLS] の特徴量を識別器の入力として用いた. 表 10 の通りにデータを分割し, 訓練データで学習しテストデータでの予測ラベルと正答率を求めた. ラベルが 1 であることを比較的高い精度で予測できているのはこの表を見る限りでは, 夕食 po, 風呂 po, サービス po, 立地 po, 部屋 po である. .

表 10: 多値分類タスクに用いたデータの内訳

総データ数	訓練データ数	テストデータ数
1200	960	240

表 11: 多値分類タスクに用いたパラメータ

値	パラメータ
入力層の次元数	768
出力層の次元数	12
バッチサイズ	12
最適化関数	Adam
損失関数	BCEWithLogitsLoss

表 10 の総データを用いて 5 分割検証を行い, 5 個のモデルを作成した. 表に各モデルでの正解率を示す. ただし, ここでの正解率とは各クラスの正解率の平均値としている. また, 図 3, 4 に最も正解率の高いモデルにおける訓練データとバリデーションデータの損失の推移を示す.

表 12: 多値分類タスクにおける 5 分割交差検証での正解率

0.8979	0.8983	0.9035	0.9031	0.8955
--------	--------	--------	--------	--------

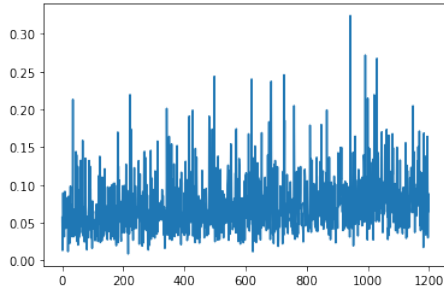


図 3: 表 12 の 3 つ目のモデルにおける訓練データの損失の推移

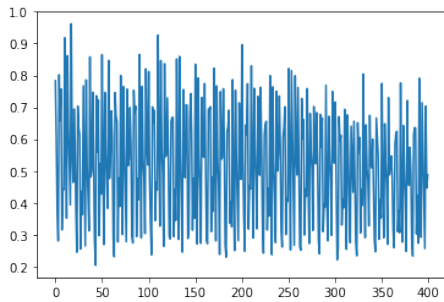


図 4: 表 12 の 3 つ目のモデルにおけるバリデーションデータの損失の推移

表 12 で示した最も正解率の高いモデルを用いて表 10 で示したテストデータに対して実験をした。そして図 5 に得られた予測データと正解データにおいて、各データにおける立っているラベル数の分布を比較したものを示す。また、この予測データにおける全ラベルが正解したものは 78 個であった。

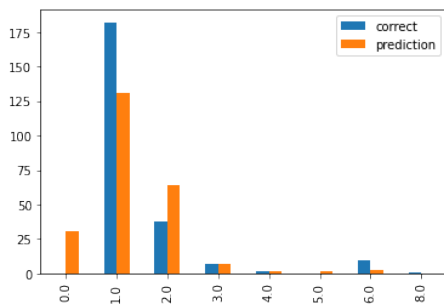


図 5: 予測データと正解データの各データにおける立っているラベル数の分布

図 5 からわかる通り、全ラベルが 0 となっている予測データが約 25 個含まれている。今回の実験をするにあたって、全ラベルが 0 となっているデータを取り除いたため、このような予測データは適切ではないと考えた。そのため、全ラベルが 0 となる予測データに関しては、各出力を 0 か 1 かに丸める際に、少なくとも 1 つのラベルが立つようにする処理を行った。図 6 に先ほどの処理を行った予測データと正解したデータの各データにおける立っているラベル数の分布を示す。

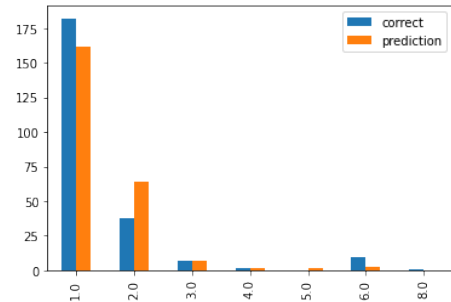


図 6: 特定の処理を加えた予測データと正解データの各データにおける立っているラベル数の分布

以降では図 6 に示した予測データを適切な予測データとして用いることとする。表に予測データの各ラベルの 1 の個数と正解データの各ラベルの 1 の個数を示す。この予測データのうち、全ラベルが正解したものは 85 個であった。全ラベルが 0 とならないようにした為、予測が正解したデータ数が増えたと考えられる。図 6 にこの予測データの各ラベルにおける立っているラベル数の分布を示す。また、表 13 に予測データにおける立っているラベルの正解について示す。表 15 に予測データにおける各ラベルでの正解率、適合率、再現率、F 1 値を示す。表 16 に予測ラベルの正誤の分類及び正答率と再現率を示す。そして表 17 には予測が正解した具体例を示す。そして表 18 には予測を間違えたデータの具体例を示す。

表 13: 予測データにおける立っているラベルの正解データ数

対象データ	データ数
全テストデータ	240
全ラベルが正解	85
一部ラベルが不正解	59
全ラベルが不正解	96

表 14: 予測データの各ラベルの 1 の個数と正解データの各ラベルの 1 の個数

	夕食 po	夕食 ne	風呂 po	風呂 ne	サービス po	サービス ne	立地 po	立地 ne	設備 po	設備 ne	部屋 po	部屋 ne
予測データ	62	6	26	3	66	35	41	1	29	25	37	16
正解データ	44	8	37	11	74	37	27	3	38	26	41	9

表 15: 予測データにおける各ラベルでの正解率, 適合率, 再現率, F 1 値

	夕食 po	夕食 ne	風呂 po	風呂 ne	サービス po	サービス ne	立地 po	立地 ne	設備 po	設備 ne	部屋 po	部屋 ne
正解率	0.8583	0.9583	0.9125	0.9583	0.8167	0.8333	0.8750	0.9833	0.8125	0.8792	0.9083	0.9208
適合率	0.8182	0.2500	0.5676	0.1818	0.6486	0.4324	0.7037	0	0.2895	0.4231	0.6829	0.3333
再現率	0.5806	0.3333	0.8076	0.6667	0.7273	0.4571	0.4634	0	0.3793	0.4400	0.7568	0.1875
F 1 値	0.6792	0.2857	0.6667	0.2857	0.4444	0.5588	0.5588	Nan	0.3284	0.4313	0.7179	0.2400

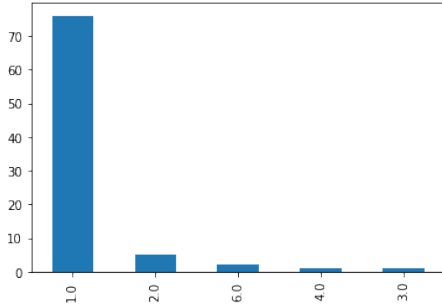


図 7: 全ラベルが正解したデータにおける立っているラベル数の分布

## 7 まとめと今後の課題

本研究では 2 値分類と多値分類をした。異なるモデルでの比較検討であったり, より適切な閾値やパラメータの設定であったり, できていないことがあった。今後の課題として, 関連研究であるアスペクトベースの感情分析に基づき, 分類された文章について文中に含まれるアスペクトカテゴリに対するフレーズの位置を特定する研究に取り組むことを考えた。

## 参考文献

- [1] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*, 2018.
- [2] Maria Pontiki, Dimitrios Galanis, Harris Papageorgiou, Suresh Manandhar, and Ion Androutsopoulos. Semeval-2015 task 12: Aspect based

sentiment analysis. In *Proceedings of the 9th international workshop on semantic evaluation (SemEval 2015)*, pp. 486–495, 2015.

- [3] 三浦義栄, 赤井龍一, 渥美雅保. 文中の複数アスペクトのセンチメント分析のための自己注意ニューラルネットワーク. 人工知能学会全国大会論文集 第 34 回全国大会 (2020), pp. 3Rin441–3Rin441. 一般社団法人 人工知能学会, 2020.

表 16: 多値分類の予測ラベルの分類

	夕食 po	夕食 ne	風呂 po	風呂 ne	サービス po	サービス ne	立地 po	立地 ne	設備 po	設備 ne	部屋 po	部屋 ne
1 を 1 と当てた	36	2	21	2	48	16	19	0	11	11	28	3
0 を 0 と当てた	170	228	198	228	148	184	191	236	184	200	190	218
1 を 0 と間違えた	8	6	16	9	26	21	8	3	27	15	13	6
0 を 1 と間違えた	26	5	32	1	18	19	22	1	18	1	9	13

表 17: 予測が正解したデータの具体例

テキスト	夕食 po	夕食 ne	風呂 po	風呂 ne	サービス po	サービス ne	立地 po	立地 ne	設備 po	設備 ne	部屋 po	部屋 ne
全体的には良かったと思います.	1	0	1	0	1	0	1	0	1	0	1	0
チェックアウトもただ鍵を返すだけなのに並ばなければならず、朝の貴重な時間が無駄になるだけ.	0	0	0	0	0	1	0	0	0	0	0	0
夕食は大変美味しかったです.	1	0	0	0	0	0	0	0	0	0	0	0
近くにコンビニや居酒屋もあり、徒歩で行ける距離です.	0	0	0	0	0	0	1	0	0	0	0	0
お部屋は、清潔で綺麗なものでゆっくり過ごすことが出来ました.	0	0	0	0	1	0	0	0	0	0	1	0

表 18: 予測を間違えたデータの具体例 (カッコ内は正解ラベルを示している.)

テキスト	夕食 po	夕食 ne	風呂 po	風呂 ne	サービス po	サービス ne	立地 po	立地 ne	設備 po	設備 ne	部屋 po	部屋 ne
立地がかなり歓楽街にあることもあり、外にでるときにわざわざ鍵を預けないといけない	0	0	0	0	0	0	1 (0)	0 (1)	0	0	0	0
2 日目は 3 日目の朝に気付いたのですが、室内電話の受話器が外されたままで、清掃の後のチェックが行き届いてないと印象が残りました..	0	0	0	0	0	0 (1)	0	0	0	0	0	1
全体的に大満足ですので、また東京に行くときはお世話になります.	1	0	1	0	1	0	0 (1)	0	1	0	0 (1)	0
売りのはずの客室露天風呂を、もう少し清潔に気持ちの良いものにして頂ければ最高です.	0	0	1 (0)	1	0 (1)	0	0	0	0	0	0	0
子供はプレゼントにとっても喜んでいました.	1 (0)	0	0	0	1	0	1 (0)	0	1 (0)	0	0	0