

カテゴリ特化型感情極性辞書を用いたホテル評価値予測のための カテゴリ分類モデルの作成

杉岡 昂輝† 楠 颯孔†† 亀井 清華†† 森本 康彦††

† 広島大学情報科学部 〒739-8521 東広島市鏡山1丁目4-1

†† 広島大学大学院先進理工系科学研究科 〒739-8521 東広島市鏡山1丁目4-1

E-mail: †{B191767,m224597,s10kamei,morimo}@hiroshima-u.ac.jp

あらまし ホテル予約サイトでは、ホテルの宿泊客から投稿された各ホテルに対する評価が提供される。各評価には、自然言語テキストによるレビューと、「サービス」や「立地」、「部屋」などの特定のカテゴリに対する数値による評価値が含まれる。柴田らの手法では、各カテゴリに関連したキーワードをもとにレビュー文に対してカテゴリを割り振り、それらをもとにカテゴリ評価値を予測しているが、キーワードによるカテゴリ分類だとすべてのレビュー文を正確に分類することは難しい。そこで、我々はBERT ファインチューニングによるカテゴリ分類モデルを作成し、キーワードによる分類ではどのカテゴリにも割り振られなかったレビュー文に対してモデルを用いることで再分類を行った。これにより、柴田らのカテゴリ特化型感情極性辞書を用いたカテゴリ評価値の予測について精度を向上することができた。

キーワード BERT, 感情極性, ロジスティック回帰, 評価値予測

1 はじめに

ホテル予約サイトである楽天トラベル[1]では、各ホテルの利用客(レビューア)から投稿された、そのホテルに対する評価が提供される。各評価は、自然言語によってホテルの印象を記述したレビューと、満足度を表す数値による評価値の組として投稿される。評価値は総合評価に加えて、「サービス」、「立地」、「部屋」、「設備・アメニティ」、「風呂」、「食事」の6つの評価視点について、それぞれ1~5の5段階評価で表される。これらの6つの評価視点を以下では**カテゴリ**と呼ぶ。これらの評価により、サイト利用者はホテル選びにおいて自らの判断基準と照らし合わせて選ぶことができ、ホテルオーナーは宿泊客の意見を知ることができる。しかし評価の中には、評価値が付与されていないものや、レビューと評価値が矛盾したものがある。これらの一部の評価は、サイト利用者を混乱させる可能性がある。このような問題を解決するため、レビュー文から評価値予測をおこなうという研究がある。それにより、評価値が付与されていないレビューに対して自動的に評価値を付与することができたり、レビューの内容と評価値が矛盾しているものが投稿されそうになった場合にレビューアに注意喚起をすることができる。

レビュー文から評価値予測を行う先行研究に、柴田らの研究[2]がある。柴田らは、レビューから各カテゴリを説明する文を抽出し、それらの文の単語をBERT[3]によってベクトル化した。BERTとは、2018年にGoogleから発表された自然言語処理モデルであり、現在でも最も注目されている自然言語処理モデルの一つである。BERTは、2種類の学習方法があり、一つは柴田らが使用したfeature-basedアプローチであり、もう一つはファインチューニングである。feature-based アプ

ローチでは、事前学習済みモデルのBERTパラメータを固定して、入力に対する特徴量を抽出する。ファインチューニングでは、個別タスクの内容に応じた分類器を事前学習済みモデルのBERTに接続し、事前学習で得られたパラメータを初期値として特徴量抽出器と分類器の両方のパラメータを学習する。そのためファインチューニングでは、比較的少数の学習データでも高い性能のモデルを得ることができる。

柴田らはfeature-basedアプローチによって得られたレビュー内の単語の平均ベクトルに対し、追加特徴量として各カテゴリで文中の単語に表れる感情の正負や強弱を数値化した感情極性値を加えることでカテゴリ別の評価値を予測し、既存の手法より高い精度を上げることができた。しかし、柴田らの研究では、レビュー文について、どのカテゴリにも割り振られない文が多く存在し、それらの文の中にもカテゴリを説明している文があると考えられる。そこで本稿では、こうした文を抽出するために、BERTファインチューニングを用いて各カテゴリごとに分類モデルを作成し、評価値予測の精度向上について有効性を示す。

2 関連研究

2.1 評価値予測の研究

まず、評価値予測の既存手法を紹介する。小林ら[4]は、ECサイトの商品レビューからBERT[3]などのニューラルネットワークを用いて商品に対する総合評価値を予測をした。しかし、彼らの手法は総合評価値のみを考慮した評価値予測であり、複数のカテゴリについての評価値予測については議論されていない。

また、楽天トラベルのホテルレビューのカテゴリ別評価値予測

についての研究には以下のものがある。まず、張ら [6] は、カテゴリに関連する文をキーワードで抽出し、それを Bag-of-Words を用いてベクトル化し、それを用いてロジスティック回帰でカテゴリ別評価値を予測した。外山ら [7] は、レビュー全体と各文を PV-DM [9] を用いてベクトル化し、ニューラルネットワークでカテゴリ別評価値を予測した。井上ら [8] は、カテゴリに関連する文をキーワードで抽出し、BERT ファインチューニングでカテゴリ別評価値を予測した。

レビューの記述内容から評価値を予測することにおいて、レビューに表れる感情を考慮することは非常に重要だと考えられるが、上記の既存手法では、レビュー中の単語に現れるポジティブな感情やネガティブな感情について明に考慮したものはない。

2.2 柴田らの研究

次に、レビューに表れる感情を考慮してカテゴリ別の評価値を予測した既存手法として、柴田ら [2] の手法を紹介する。

まず柴田らは、田熊らが作成したカテゴリ辞書 [5] を基に、ある単語がレビューに含まれている場合にそのレビューに記述されているカテゴリを特定できるような単語を手で抽出し、**カテゴリ辞書**を作成した。カテゴリ辞書の例を表 1 に、登録単語の内訳を表 2 に示す。そして各レビューを文に分割し、カテゴリ辞書を用いて各文が言及しているカテゴリのラベルを付与する。さらに、各レビューの各カテゴリに該当する文を抽出することで**カテゴリレビュー**を作成した。このとき、各カテゴリに対してランダムに 20 件ずつカテゴリレビューを抽出してアンケートを行い、回答結果から 90.83% のカテゴリレビューに対するラベル付けが妥当であると判断された。

また、柴田らはカテゴリごとに感情を表す単語 (感情極性語) とその語が持つ感情の正負と強弱を数値化した感情極性値の組を要素とする**カテゴリ特化型感情極性辞書 (CSPD: Category-oriented Sentiment Polarity Dictionary)**を作成した。CSPD の例を表 3 に示す。この辞書には、各カテゴリにおよそ 3,000~4,000 程度の感情極性語が登録されており、極性値は $-7.80674 \sim 6.37931$ となっている。この極性値は、楽天トラベルデータセットにおける、各カテゴリと各評価値での単語の出現率に基づいて計算され、平均値が 0 に近くなるように補正されたものとなっている。

彼らは、カテゴリレビューの単語ベクトル化手法として BERT を用い、カテゴリレビュー内のすべての単語の平均ベクトルを用いた。さらに CSPD を用いた感情極性値を特徴量として加えたベクトルをロジスティック回帰に入力することで、カテゴリ別評価値を予測した。この手法を用いることで BERT 単体によるカテゴリ別評価値予測よりも高い精度で予測ができていた。

しかし、柴田らのカテゴリ辞書では分割した文に対して 35% 以上がどのカテゴリにも割り振られず、これらの文の中にもカテゴリについて言及した文が含まれているものと考えられる。

表 1 カテゴリ辞書の例.

サービス	立地	部屋	設備	風呂	食事
おもてなし	コンビニ	エアコン	PC	サウナ	おにぎり
お金	タクシー	カーテン	エレベータ	シャワー	ごはん
⋮	⋮	⋮	⋮	⋮	⋮

表 2 カテゴリ辞書の内訳.

サービス	立地	部屋	設備	風呂	食事	総語彙数
74	50	97	89	30	56	330

表 3 カテゴリ特化型感情極性辞書の例.

	サービス	立地	部屋	設備	風呂	食事
ひどい	-6.1	-6.0	-7.0	-6.7	-6.5	-7.2
素晴らしい	4.5	極性値無し	5.2	5.3	5.7	4.5
⋮	⋮	⋮	⋮	⋮	⋮	⋮

表 4 BERT ファインチューニングのパラメータ

	パラメータ
学習率	$1e-5$
epoch 数	3
バッチサイズ	16

3 提案手法

本研究では、柴田らの研究 [2] においてカテゴリ辞書でラベルが付かなかった文について、カテゴリ分類モデルを作成して再分類を行い、カテゴリレビューの追加、更新を行う手法を提案する。

柴田らのカテゴリ辞書によって各カテゴリに割り振られた文を教師データとし、各文に各カテゴリのラベルが付くか否かを分類するように、日本語 BERT 訓練済みモデル [10] を用いてファインチューニングを行う。このとき、教師データについては偏りがないように各カテゴリで正解ラベルと不正解ラベルのデータ数が 2,500 件ずつと同じになるように調整をし、カテゴリ毎に二値分類を行うカテゴリ分類モデルを作成した。これは各入力文について複数のカテゴリラベルが付与されることを許すようにするためである。そして、完成した 6 つのカテゴリ分類モデルを用いて、柴田らのカテゴリ辞書でラベルが割り振られなかった文を再分類した。

4 実験評価

4.1 使用データセット

本実験では、2014–2019 年の楽天トラベルデータセット [11] を用いた。また、2014–2016 年のデータセットを**旧データ**、2017–2019 年のデータセットを**新データ**と分け、カテゴリ別評価値予測を行う際に旧データを訓練データ、新データをテストデータとして使用した。また、カテゴリ分類モデルを構築する際に使用した文は、2016 年のレビューをカテゴリ辞書を用いて分類した上で、各カテゴリについてランダムに採取した。

旧データのレビュー数は 1,173,803 件でこれを文に分割す

ると、5,376,801 文となり、カテゴリ辞書のみではそのうちの 36.1% (1,941,119 文) がどのカテゴリにも割り振られなかった。

新データのレビュー数は 1,497,595 件でこれを文に分割すると 6,382,940 文となり、カテゴリ辞書のみではそのうちの 35.6% (2,274,380 文) がどのカテゴリにも割り振られなかった。

4.2 カテゴリ分類モデルによる再分類

BERT ファインチューニングによるカテゴリ分類モデルを用いて、カテゴリ辞書で分類されなかった文 (旧データ:1,941,119 文, 新データ:2,274,380 文) に対してカテゴリの再分類を行った。このとき、各パラメータは表 4 のとおりとした。結果を表 5 に示す。なお、一つの文に対して複数のカテゴリが割り振られることがあるので、表中の「全体」は各カテゴリの「分類できた文の数」の合計にはならないことに注意されたい。分類率は (分類モデルで分類できた文の数)/(カテゴリ辞書で分類されなかった文の数) を表している。結果から「風呂」以外のカテゴリにおいてはカテゴリ辞書で分類されなかった文の中から 2~3%の文を各カテゴリに再分類することができた。また、全体としては 11%前後の文を各カテゴリに分類することができた。

以降の実験については、旧データ、新データのカテゴリレビューを以下のように type A~C に分類して実験を行った。各 type のデータ数は表 6 のとおりである。

- *type A*: カテゴリ辞書で分類されていたカテゴリレビューに対し、カテゴリ分類モデルによって新たな文が追加されたカテゴリレビュー。
- *type B*: カテゴリ辞書で分類されているカテゴリレビューに対し、カテゴリ分類モデルによって新たな文が追加されなかったカテゴリレビュー。
- *type C*: カテゴリ分類モデルで新たに作成された、カテゴリ辞書のキーワードを含まないカテゴリレビュー。

4.3 カテゴリレビューの妥当性

まず、被験者へのアンケートによって各カテゴリに再分類された文の妥当性を評価した。ここでは、カテゴリ分類モデルによって新たに作成された type C の各カテゴリレビューが、そのカテゴリを説明しているか否かを評価した。旧データから各カテゴリで type C のカテゴリレビューとなったものからランダムに 20 件ずつ抽出し、カテゴリが偏らないようにシャッフルをした上で、各カテゴリレビューに対して 3 人の被験者を割り当てた。被験者は 10 代 1 名 (男性: 1 名), 20 代 17 名 (男性: 15 名, 女性: 2 名) の計 18 名で行った。各被験者は与えられた文を読み、各文に記載されている内容のカテゴリを選択する。このとき、各文に対して複数のカテゴリを選択しても良いものとし、また、一つもカテゴリを選択しないことも可能であった。被験者に対する説明は以下のとおりである。

「次のホテルレビューの一部を読み、そのレビューが記述していると思うカテゴリ (“サービス”, “立地”, “部屋”, “設備・アメニティ”, “風呂”, “食事”) を選択してください。カテゴリは複数選択可能です。また、一つも選ばないという選択も可能です。」

カテゴリ X のカテゴリレビュー r_X に対して、被験者が選択したカテゴリに X が含まれていた場合に、その被験者は正解したものととした。 r_X に対して、3 人の被験者のうちの少なくとも 1 人が正解した場合に、 r_X は X について妥当であるとした。なお、ある文 r が複数のカテゴリのラベルを有していた場合は、そのそれぞれのカテゴリで判断した。アンケートの集計結果を表 7 に、3 人ともが正解したレビューの例を表 8 に示す。

アンケート結果から、分類が妥当とされたカテゴリレビューの割合は全体の 71.3%となった。カテゴリごとに見ていくと、「サービス」、「立地」、「風呂」、「食事」は正解率が 76%を超えており、妥当性は十分に高いことがわかるが、「部屋」、「設備」については正解率が 50%台とうまく分類できていなかった。また、表 8 を見ると、どのレビューも各カテゴリについて説明している文だとわかり、カテゴリ辞書では分類されなかった文について上手く分類できていることがわかる。

4.4 カテゴリ評価値予測

ここでは、提案手法によるカテゴリレビューを用いた場合 (dic+model) と、柴田らの (カテゴリ辞書のみによる) カテゴリレビューを用いた場合 (dic) とで、カテゴリ評価値予測を行い、精度について比較を行う。カテゴリ評価値予測は、柴田らの手法に従い、BERT によるカテゴリレビューの平均単語ベクトルに CSPD を用いた感情極性値を加えたベクトルを用い、ロジスティック回帰によって各評価値に分類することで行う。正解のカテゴリ評価値は、レビューアーによるカテゴリ評価値とした。さらに、感情極性値の影響も確認するために、それぞれ感情極性を加えなかった場合のカテゴリ評価値予測の精度も比較する。

4.4.1 データセット

本実験で使用するデータセットについて説明する。

まず、評価値予測のための分類モデル (以下、**評価値予測モデル**と呼ぶ) の作成では、旧データから各カテゴリについて 100,000 件を訓練データとして用いた。ただし、カテゴリレビューは type A をすべて用い、足りない数を type B のデータからランダムに抽出した。このとき dic については、type A のカテゴリレビューのうち、カテゴリ辞書のみで抽出された部分だけを用いることに注意されたい。そして、各カテゴリについて 5 回訓練データの抽出をし、そのそれぞれで評価値予測モデルを作成した。

また、テストデータは新データから各カテゴリについて 100,000 件を用いた。ただし、この場合も、type A のデータすべてを用い、足りない数を type B のデータからランダムに抽出した。このとき、各カテゴリで 5 つのモデルに対しては共通のテストデータを用いた。

4.4.2 評価値予測モデル

評価値予測モデルとして、scikit-learn のロジスティック回帰 [12] を用いた。各カテゴリ X について、日本語 BERT 訓練済みモデル [10] を用いて、各カテゴリレビュー r_X のすべての単語をそれぞれ 1024 次元でベクトル化して平均をとり、 r_X のベクトルとする。さらに、 X のための CSPD を基に、 r_X に含

表5 カテゴリ分類結果

データ		サービス	立地	部屋	設備	風呂	食事	全体
旧	分類できた文の数	50,340	60,295	73,055	41,122	6,248	35,604	226,949
	分類率 (%)	2.6	3.1	3.8	2.1	0.3	1.8	11.7
新	分類できた文の数	55,376	69,334	60,191	51,567	7,645	43,281	248,424
	分類率 (%)	2.4	3.0	2.6	2.3	0.3	1.6	10.9

表6 タイプ別のカテゴリレビュー数

type	データ	評価値	1	2	3	4	5
A	旧	サービス	4,180	2,838	4,937	7,247	13,312
		立地	225	1,018	3,768	10,350	15,649
		部屋	3,756	5,196	9,268	15,966	15,441
		設備	1,917	2,990	6,119	7,052	5,596
		風呂	221	337	521	1,230	1,739
		食事	1,294	1,664	2,627	6,676	13,208
	新	サービス	5,432	3,434	5,163	7,257	13,588
		立地	278	1,213	4,134	11,194	17,714
		部屋	4,095	4,960	7,589	12,761	12,379
		設備	2,668	3,716	7,426	8,429	7,266
		風呂	334	404	610	1,528	2,066
		食事	1,814	2,032	3,005	7,594	15,393
B	旧	サービス	18,629	26,390	113,319	197,930	267,705
		立地	2,558	14,839	56,017	161,985	236,504
		部屋	20,211	41,205	104,791	209,965	239,538
		設備	16,067	37,528	115,183	144,811	117,338
		風呂	17,972	32,467	56,452	125,764	183,333
		食事	14,128	27,673	61,072	152,826	253,147
	新	サービス	26,224	33,643	131,185	229,382	321,737
		立地	3,242	17,877	67,613	190,682	289,171
		部屋	30,403	54,154	127,218	258,779	306,701
		設備	22,373	47,002	140,881	174,911	149,520
		風呂	24,702	41,565	70,593	157,891	232,957
		食事	19,109	34,338	73,698	189,539	317,603
C	旧	サービス	408	530	2,562	4,022	5,578
		立地	134	675	3,218	8,599	10,956
		部屋	353	908	3,442	5,639	4,514
		設備	483	1,115	3,676	4,620	3,808
		風呂	65	137	384	607	608
		食事	59	158	687	1,559	1,746
	新	サービス	614	664	3,080	4,644	6,576
		立地	191	856	3,930	10,601	13,111
		部屋	440	884	2,884	4,592	3,713
		設備	654	1,465	4,743	5,883	4,573
		風呂	86	186	471	708	834
		食事	97	220	856	2,148	2,411

表7 アンケート結果

正解者数	サービス	立地	部屋	設備	風呂	食事	合計
0人	2	4	20	14	5	3	48
1人	2	5	9	9	3	2	30
2人	8	3	5	2	5	3	26
3人	12	13	8	9	8	13	63
合計	24	25	42	34	21	21	167
正解率 (%)	91.7	84.0	52.4	58.8	76.2	85.7	71.3

まれるすべての感情極性語の感情極性値の平均値を追加特徴量として加え、1025次元のベクトルを作成する。これを、評価値予測モデルへの入力とする。

表8 アンケートで妥当とされたカテゴリレビューの例

サービス	無料のシャトルバスは大変ありがたかったです
立地	市街地から少しだけ離れて、静かな場所です
部屋	壁が薄くてまわりの音が気になりました
設備	ヘアブラシの形状が悪く使い難いので改善求む
風呂	岩盤浴も楽しむことができ、飽きずにのんびり過ごせました
食事	ソフトドリンク飲み放題・朝カレーもうれしい

4.4.3 評価指標

各手法でそれぞれ作成した5つの各評価値予測モデルで予測した結果は、以下の5つの指標で評価する。レビュー者による評価値(正解)を Y 、評価値予測モデルによる予測値を Y_e とすると、

- 正答率: すべてのレビューのうち、 $Y = Y_e$ となった割合。
- 適合率: Y_e となったレビューのうち、 $Y = Y_e$ となった割合。
- 再現率: Y のレビューのうち、 $Y = Y_e$ となった割合。
- F1値: 適合率と再現率の調和平均値。
- RMSE: Y と Y_e の二乗平均平方根誤差。

なお、適合率、再現率、F1値は、評価値ごとに計算し、その平均値を用いる。また、以下では5つの評価値予測モデルによる各指標の平均値を用いて比較する。

4.4.4 結果

各カテゴリの評価値を予測した結果を表9に示す。各結果についてt検定を行い、表中の†は有意水準1%で、*は有意水準5%でそれぞれdic+modelと比較をした際に統計的に有意な差が示されたことを表す。また、4つの手法のうち最も良い性能を示した場合を太字で記載している。

表9より、「風呂」以外では柴田ら[2]のCSPDによる感情極性とカテゴリ分類モデルを用いた方が予測精度が一番高くなった。一方で「風呂」は有意な差はほとんど見られなかった。これは、表6のtype Aの新旧データについてカテゴリレビュー数が少なく、特に評価値1~3のカテゴリレビュー数が極端に少ないことから、十分に学習できていないのが原因だと考えられる。また、感情極性なしの場合で分類方法による精度を比べてみると、感情極性がある場合と同様にカテゴリ分類モデルを用いたほうが予測精度が高くなった。このことから、評価値予測モデルへの入力として、カテゴリを説明している文の情報を適切に増やすことで、予測精度が高くなることが分かった。

よって、提案手法では、柴田らのカテゴリ辞書で割り振られなかったレビューについてBERTファインチューニングによる再分類を行うことで、カテゴリ別評価値予測について精度の向上を達成することができた。

表9 カテゴリ評価値予測結果

	分類方法	感情極性	サービス	立地	部屋	設備	風呂	食事	平均
正答率	dic	有	0.5261†	0.5388†	0.4748†	0.4220†	0.5254	0.5750†	0.5103†
	dic+model	有	0.5377	0.5440	0.4821	0.4278	0.5253	0.5833	0.5167
	dic	無	0.5255†	0.5381†	0.4729†	0.4207†	0.5245	0.5730†	0.5091†
	dic+model	無	0.5369	0.5424*	0.4786†	0.4264	0.5250	0.5807*	0.5150†
適合率	dic	有	0.4491†	0.3539†	0.4283†	0.4004†	0.4283	0.4475†	0.4179†
	dic+model	有	0.4602	0.3638	0.4386	0.4083	0.4293	0.4579	0.4264
	dic	無	0.4476†	0.3538†	0.4253†	0.3989†	0.4276	0.4442†	0.4162†
	dic+model	無	0.4591	0.3598	0.4294†	0.4079	0.4313	0.4540*	0.4236†
再現率	dic	有	0.4391†	0.2811†	0.4089†	0.3582†	0.4054	0.4137†	0.3844†
	dic+model	有	0.4551	0.2891	0.4206	0.3696	0.4034	0.4262	0.3940
	dic	無	0.4382†	0.2789†	0.4060†	0.3566†	0.4047	0.4104†	0.3825†
	dic+model	無	0.4548	0.2866	0.4189*	0.3689	0.4022	0.4245	0.3927†
F1 値	dic	有	0.4360†	0.2905†	0.4139†	0.3688†	0.4116	0.4247†	0.3909†
	dic+model	有	0.4503	0.3000	0.4253	0.3800	0.4108	0.4360	0.4004
	dic	無	0.4347†	0.2878†	0.4107†	0.3672†	0.4108	0.4213†	0.3888†
	dic+model	無	0.4500	0.2969	0.4193†	0.3797	0.4106	0.4335*	0.3987†
RMSE	dic	有	1.0314†	0.9031†	1.0327†	1.0647†	1.0008†	0.9450†	0.9963†
	dic+model	有	0.9862	0.8905	1.0079	1.0494	1.0126	0.9175	0.9774
	dic	無	1.0353†	0.9067†	1.0420†	1.0690†	1.0055†	0.9560†	1.0024†
	dic+model	無	0.9912†	0.8959†	1.0229	1.0540†	1.0030	0.9257†	0.9821†

5 おわりに

本稿では、ホテルのレビューから各文を各カテゴリに分類する方法を提案した。具体的には、各カテゴリにおいて、BERT ファインチューニングを用いてカテゴリ分類モデルを作成した。提案したカテゴリ分類モデルを用いることで、「風呂」以外のカテゴリでは、柴田らの手法による評価値予測においても優れた結果を残すことができた。

アンケートの結果から、「部屋」、「設備」のカテゴリについては分類の精度が低いことが分かったので、今後の課題の一つとして、これらのカテゴリに関するカテゴリ分類モデルの作成について手法の洗練が必要である。また、本稿ではカテゴリ評価値予測の手法として柴田らの手法を用いたが、レビューデータをカテゴリ分類した上で評価値を予測する他の手法においても同様に精度を向上することが出来ているかどうかを確認する必要がある。さらに、CSPD による感情極性値を用いるために BERT による単語の平均ベクトルとロジスティック回帰を用いたが、BERT ファインチューニングなどの深層学習による評価値予測への感情極性値の応用も検討する。

謝 辞

本研究では、国立情報学研究所の IDR データセット提供サービスにより楽天グループ株式会社から提供を受けた「楽天データセット」(https://rit.rakuten.com/data_release/)を利用した。

文 献

- [1] 楽天グループ株式会社: 楽天トラベル, <<https://travel.rakuten.com/>>.
- [2] 柴田 諒人, 亀井 清華, 中野 浩嗣, “ホテルの評価値予測のためのカテゴリ特化型感情極性辞書,” 情報処理学会論文誌データベー

ス, Vol.14, No.3, pp. 16–29, (2021).

- [3] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, “BERT: pre-training of deep bidirectional transformers for language understanding”, In Proc. Annual Conference of the North American Chapter of the Association for Computational Linguistics, pp. 4171–4186, (2019).
- [4] 小林 義幸, 越仲 孝文, “EC サイトのレビューテキストからのレーティング予測と購買者評価の分析,” 2022 年度人工知能学会全国大会 (第 36 回), (2022)
- [5] K. Takuma, J. Yamamoto, S. Kamei and S. Fujita, “A Hotel Recommendation System Based on Reviews: What Do You Attach Importance To?”, In Proc. Fourth International Symposium on Computing and Networking, pp. 710–712, (2016).
- [6] 張 博, 白井 清昭, “レビューテキストの書き手の評価視点に対する評価点の推定,” 言語処理学会第 23 回年次大会発表論文集, pp.803–806, (2017).
- [7] 外山 洋太, 三輪 誠, 佐々木 裕, “文書・文間及びカテゴリ間の関係を考慮したレーティング予測,” 言語処理学会第 22 回年次大会発表論文集, pp.158–161, (2016).
- [8] 井上 敬通, 藤田 桂英, “旅行サイトにおけるレビュー文のカテゴリを考慮した評価値予測,” 情報処理学会第 82 回全国大会講演論文集, pp.467–468, (2020).
- [9] Q. Le and T. Mikolov, “Distributed representations of sentences and documents,” In Proc. 31st International Conference on International Conference on Machine Learning, pp.II–1188–II–1196, (2014).
- [10] 東北大学 乾研究室. Pretrained Japanese BERT models. <https://github.com/cl-tohoku/bert-japanese>
- [11] 楽天グループ株式会社: 楽天データセット, 国立情報学研究所情報学研究データリポジトリ (データセット), DOI: 10.32130/idr.2.0, (2014).
- [12] D.R. Cox, “The regression analysis of binary sequences,” Journal of the Royal Statistical Society: Series B (Methodological), Vol.20, No.2, pp.215–232, (1958).