

進捗報告

1 はじめに

2 要素技術

2.1 BERT

3

BERT-[1] は があります。

4 データセット

4.1 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。

4.2 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

5 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、制度評価を行った。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。実験の評価指標として

5.1 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 値分類を行った。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データの内訳

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

5.2 実験 2

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.

表 7: 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	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

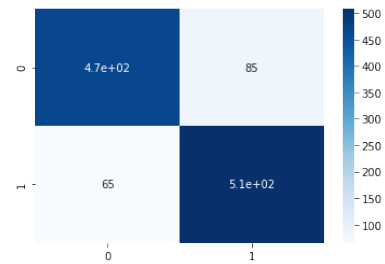


図 1: 5 分割検証を行った場合のコンフュージョンマトリクス

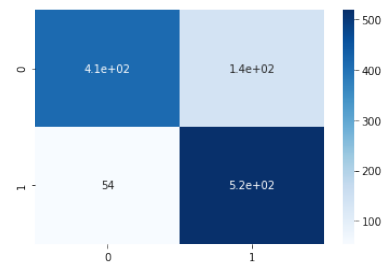


図 2: 訓練データをそのまま用いた場合でのコンフュージョンマトリクス

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
<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식을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식을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んですね。</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、チャラチャラしていない老舗ホテルの貫禄があります。</p>	1	0