4C-06

BERT による日本語文の感情分析と話題分析

圓谷顯信† 高橋宏和† 安達由洋†
東洋大学総合情報学部総合情報学科†

1. はじめに

BERT[1]は文の分散表現をコンテキストに依存して動的に生成するので、話題や感情による文の分類・検索の精度が Word2Vec より上がることが期待される。BERT による日本語文のクラス分類に関する研究やセンチメントアナリシスに関する研究は幾つか報告されているが、細粒度の感情分析に関する研究は見当たらない。一方、我々は感情語を細粒度（11 感情カテゴリ）に分類した感情語辞書を用いて、比較的精度よく高速に感情分析するシステム EEAS[2]を報告している。また、Word2Vec により日本語 Wikipedia 全文から学習した日本語 Wikipedia エンティティベクトル[3]を用いて日本語テキストを分類・検索する研究[4]も報告している。

本稿では、BERT を用いて文献[2]と同じ 11 感情カテゴリからなる日本語文の感情分析を行った結果を報告する。そして、実験結果をもとに EEAS と BERT の長所・短所を明らかにし、それぞれが適した応用分野について論じる。また、BERT を用いた話題に基づく日本文の分析手法についても報告する。特に、BERT を特徴抽出器として用い、得られた特徴ベクトルを用いて日本語文データセットをクラスタリングする研究についても報告する。

2. 感情に基づく日本語文の分析

日本語文を感情分析する基本となる感情分類法として 11 感情カテゴリベクトル［喜,怒,哀,怖,恥,好,厭,昂,安,驚,望］[5]を用いる。BERT の事前学習済みモデルとして、東北大学の Pretrained Japanese BERT models（東北大学モデルと呼ぶ） と、NICT の BERT 日本語 Pre-trained モデル BPE なし（NICT モデルと呼ぶ）を採用した。感情分析対象として、Amazon レビューから収集した 6,834 文、Google クチコミから収集した 3,150 文、TSUKUBA コーパスから収集した 3,132 文の合計 13,116 文に対して著者等 3 人の多数決で教師ラベルを付けた日本語文データセットを用意した。そして、この日本語文データセットをシャッフルしたのち、訓練データ 7869 文、検証データ 2623 文、テストデータ 2624 文に分割した。

本稿では、東北大学モデルを使用した実験結果をまとめる。

2.1 感情に基づく分析精度

日本語文データセットの訓練データと検証データを用いて東北大学モデルを fine-tuning し、11 感情分析 BERT モデルを作成した。表 1 に、このモデルによるテストデータに対する精度評価結果を示す。なお、表 1 の中で文数とは各感情カテゴリに属する文の数を表す。

表 1． BERT による 11 感情分析精度

	文数	accuracy	precision	recall	F1score
喜	462	0.919	0.822	0.690	0.751
怒	16	0.995	0.714	0.313	0.435
哀	24	0.989	0.417	0.417	0.417
怖	9	0.998	0.75	0.667	0.706
恥	1	1.00	0.00	0.00	0.00
好	321	0.926	0.752	0.586	0.658
厭	188	0.942	0.586	0.654	0.618
昂	14	0.994	0.00	0.00	0.00
安	25	0.994	0.786	0.44	0.564
驚	23	0.997	0.792	0.826	0.809
望	88	0.983	0.824	0.636	0.718

表 1 に示されているように、‘喜’、‘好’、‘厭’、‘望’の感情は比較的精度よく抽出ができている。

11 感情分析 BERT モデルと EEAS の感情分析精度を比較すると、BERT の方が約 10%以上 F1Score の高い感情カテゴリが幾つかある。EEAS は授業アンケート回答文や各種レビュー文に含まれる感情を極力検出するように再現率(recall)を高く調整してある。再現率を落として適合率(precision)を上げるように調整すれば EEAS の F1Score をもっと上げることができると思われる。

2.2 感情に基づく分析速度

11 感情分析 BERT モデルは multi-label text classification モデルでなく、感情カテゴリごとに sentence classification モデルで学習している。表 2 に fine-tuning および予測時間の測定結果を示す。ただし、実験環境は CPU：Intel Core i7 4790K (4 cores, 4.00 GHz)、RAM：16GB、GPU：NVIDIA GeForce RTX 3060 Ti (8 GB)である。

表 2. fine-tuning および予測時間（秒）

	GPU 使用	CPU のみ
fine-tuning 時間／epoch	140.0	－
予測時間（for ループ）	0.01352	0.19809
予測時間（mini-batch）	0.00215	－

Emotion analysis and topic analysis of Japanese sentences by BERT
† Kenshin TSUMURAYA, Hirokazu TAKAHASHI, Yoshihiro ADACHI・Toyo University

この表2から、10,000文の感情分析の実行に、GPUマシンでmini-batchで実行して21秒かかることが分かる。CPUのみのコンピュータの場合、数分以上かかることが予想される。ちなみに、EEASは10,000文の感情分析をCPUのみで約1秒で実行する。

3. 話題に基づく日本語文の分類・抽出

授業アンケート回答文や各種レビュー文のクラス分類モデルを作成するとき、どのようなクラスの集合を選択するかは重要で難しい問題である。また、そのクラスの集合に対応する教師ラベル付きデータセットの収集は大変な作業となる。

そこで本研究では、BERTは日本語文の特徴抽出のみに利用し、その特徴を用いてクラスタリングする実験を行った。この実験内容は文献[4]で報告した日本語文の分類・抽出技法の、文の分散表現を得る道具を日本語Wikipediaエンティティベクトルから日本語BERT東北大学モデルに変えたものである。

3.1 BERTによる特徴抽出とクラスタリング

実験手順は以下のようである：

手順A1：東北大学モデルに日本語文データセットを入力して各文に対する768次元の分散表現を出力する

手順A2：分散表現を線形時間アルゴリズムであるk-means++法でクラスタリングする

手順A3：各クラスタに対して、適切なクラスタラベルを付ける

本実験では、｛病気,ゲーム,小説,服,勉強,大学,被害,動物,料理,政治,スポーツ,天気｝の12クラスに容易に分けることができる150文から成るクラスタリング用テストデータを使用した。図1に、クラスタリングした結果の一部を示す。

```
[('昨日のラグビーの試合で、日本がアイルランドに勝利した。', 0.95082647),
('彼らのチームの甲子園出場が決定した。', 0.93991965),
('我がチームのコーチとしてプロのラグビー選手がやって来た。', 0.93000495)]

[('近所の定食屋の蕎麦がおいしい', 0.9513715),
('麻婆豆腐をご飯にかけて食べるのが好きだ。', 0.9461688),
('吉野家の牛丼は美味しい。', 0.9459629)]

[('拓哉は白いＴシャツを着てジーパンを履いています。', 0.9697617),
('花子は赤いワンピースを着ています。', 0.96836007),
('裕太は緑色のジャケットを羽織っています。', 0.9625381)]

[('理系の学部の生徒は、埼玉県のキャンパスに通うことになる。', 0.93984926),
('彼は埼玉大学の大学院生です。', 0.9335534),
('統計学の講義は難しい。', 0.9318895)]

[('両側の唾液腺に腫脹が見られ，流行性耳下腺炎の疑いを認める。', 0.95578027),
('両方の耳下腺に腫れが見られ，おたふくかぜではないかと疑う。', 0.95424247),
('定期検査で、脳に腫瘍が認められた。', 0.95211834)]

[('一輝はテニスをやっていて、ダブルスが得意です。', 0.95190114),
('澤野はバレー部のキャプテンで、サーブが得意だ', 0.9451197),
('彼女は卓球歴が長く、大会で優勝したこともある。', 0.9436895)]
```

図1. BERT分散表現を用いたクラスタリング結果

3.2 類似文の検索

実験手順は以下のようである：

手順B1：東北大学モデルに日本語文データセットを入力して各文に対する768次元の分散表現を出力する

手順B2：問合せ文の分散表現を東北大学モデルで求める

手順B3：日本語データセットの分散表現から問合せ文の分散表現とコサイン類似度の高い文を抽出する

図2に、BERT分散表現を用いた類似文の抽出結果を示す。この類似文の検索ではノイズが多く、文献[4]で発表した日本語Wikipediaエンティティベクトルに基づく類似文検索のほうが高速で精度が高い。

```
検索文：'ゲームについて'
[('最近はスマホのゲームばかりプレイしている。', 0.7745116), ('プログラミングの講義
です', 0.75849646), ('中古のゲームを買って遊ぶことが多いです。', 0.75572795), (

検索文：'動物について'
[('イルカとクジラは、生物分類上で同じ生き物です。', 0.7589848), ('この経済学の講義
面白い', 0.74611557), ('統計学の講義は難しい。', 0.74368143), ('ハムスターをペッ

検索文：'小説について'
[('江戸川乱歩の小説は面白い', 0.7874391), ('統計学の講義は難しい。', 0.77751905
の経済学の講義はよく理解できた。', 0.772808), ('もうすぐ基礎数学のテストがある。'

検索文：'勉強について'
[('この経済学の講義はよく理解できた。', 0.79651684), ('プログラミングの講義が好き
84), ('プログラミングの講義が嫌いです', 0.77679044), ('もうすぐ基礎数学のテストが
```

図2. BERT分散表現を用いた類似文の検索結果

５．まとめ

BERTを用いた日本語文の感情分析と話題による分類・検索に関する研究を行った。11感情分析BERTモデルによる感情分析は、計算時間がかかるが幾つかの感情カテゴリでEEASよりF1Scoreが10%以上高い。授業中のアンケート文や会議中の参加者アンケート文の分析などリアルタイム処理が必要なときはEEASを使用し、時間をかけることができるSNSや各種レビューの分析には精度の高いBERTモデルを使用すればよい。話題に基づく日本語文の分類・検索についてはBERTにより得た分散表現により、クラスタリングおよび類似文の抽出が可能であることを検証した。

今後の課題として、より大規模な教師ラベル付き日本語文データセットを作成して、精度の高い11感情分析BERTモデルを開発することである。また、BERTで抽出したクラスタリングでは、適切なクラスタラベルの付加方法の検討が必要である。

参考文献

[1] Jacob Devlin, Ming-Wei Chang, Kenton Lee, Kristina Toutanova, “BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding”, arXiv:1810.04805v2 (2018).

[2] 安達由洋，近藤友啓，小林孝充，惠谷菜央，石井解人，“感情語辞書を用いた日本語文の感情分析”，可視化情報 Vol.41 No.161 (2021).

[3] 東北大学，“日本語Wikipediaエンティティベクトル”，http://www.cl.ecei.tohoku.ac.jp/~m-suzuki/jawiki_vector/ (2022/01/05アクセス).

[4] Yoshihiro Adachi and Takanori Negishi, “Development and evaluation of a real-time analysis method for free-description questionnaire responses”, Proc. the 15th IEEE International Conference on Computer Science and Education (August, 2020).

[5] 瀬山透矢，Amilcare Astremo，安達由洋，“ソーシャルメディアおよびECサイトでのレビュー分析のための‘望’感情の抽出”，FIT2021 (2021).