

アスペクトベース評判分析モデルの提案と有効性の予備的検証

1 はじめに

昨今の人工知能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を示すことが難しいという問題点がある。そこで、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。このような研究に取り組むことを目標として、本研究では新たに BERT と Transformer エンコーダを接続したモデルの提案を試みる。このモデルは分類タスクをした時に最も注目されたトークンの抽出を目的として、関連研究 [1] をベースに改良に取り組んだモデルである。

2 要素技術

2.1 BERT

BERT[2] 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によりさまざま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¹を使用した。

2.2 Transformer

Transformer[3] は 2017 年に発表された深層学習モデルである。発表以前までは、翻訳などの入力文章を別の文章で出力するというモデルは Attention を用いたエンコーダ、デコーダ形式の再帰的ニューラルネットワーク (Recurrent Neural Network, RNN) や畳み込みニューラルネットワーク (Convolution Neural Network, CNN) が主流であった。しかし、当モデルは再帰や畳み込みは行わず、Attention のみを

用いていることから並列化も可能であるため、訓練時間の削減が実現された。具体的には自己注意機構 (self-attention) が RNN や CNN と大きく異なる特徴であり、読み込んだデータの距離に関わらず、どのデータの値を読み込むかどうかを制御することができる。また、この Transformer のエンコーダ層は BERT に用いられている。本研究では Transformer のエンコーダ層を BERT のエンコーダ層に接続して用いる。具体的なモデルについては後述する。

3 関連研究

3.1 アスペクトベースの感情分析

感情分析はテキストから意見や感情、態度を分析することを目的とするタスクであり、あるコンテンツの極性がポジティブ (肯定的) かネガティブ (否定的) かを推定して分類する。自然言語処理技術の発展にともない、SNS 上での書き込みやネット上のレビューサイトにおいて、ユーザーの意見を定量的に評価することの需要が高まっている。また、感情分析も注目されている。感情分析のタスクに、文脈情報の解析に重点を置いたアスペクトベースの感情分析 [1] がある。アスペクト情報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言及しているかを示すエンティティと、エンティティが対象の側面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。文章中に含まれるアスペクト情報を利用して、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。一般に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する。最初に与えられたアスペクトカテゴリに文章を分類する。次に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。最後にフレーズの極性を分析する。そして全体としてその分析精度を向上させることを目指す。

¹<https://www.nlp.ecei.tohoku.ac.jp/news-release/3284/>

4 提案手法

4.1 BERT-Transformer モデル

アスペクトベースの感情分析において渥美らが提案したモデル [1] を参考に BERT エンコーダ層に Transformer エンコーダ層を接続した BERT-Transformer モデルを作成した。BERT エンコーダ層に Transformer エンコーダ層を接続した目的は、Transformer エンコーダ層でアスペクトカテゴリと各トークンの関連度を計算することや、Attention の可視化をすることである。図 1 に BERT-Transformer モデルの概要図を示す。

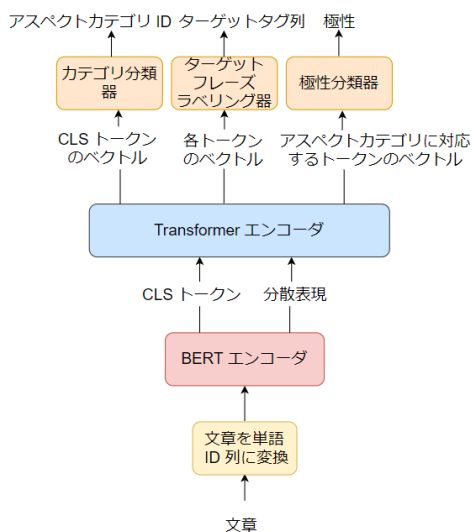


図 1: BERT-Transformer モデル概要図

まず、文章を東北大学の乾・鈴木研究室によって公開されている BERT 日本語 Pretrained モデルを用いてトークン情報を獲得し、BERT エンコーダ層に入力する。次に、BERT エンコーダ層の出力として文章データの 1 単語目 “[CLS]” と長さ 512 のトークンの特徴量を得る。それらを Transformer エンコーダ層に入力し、Attention を計算した結果をカテゴリ分類器、ターゲットフレーズラベリング器、極性分類器に入力する。カテゴリ分類器では入力文章がどのラベルに属するかを、1 つ目のトークンの “[CLS]” を用いて分類する。ターゲットフレーズラベリング器ではそれぞれの単語トークンがターゲットフレーズ（アスペクトカテゴリと類似度が高いと推定された単語）か否かの 2 値分類をする。極性分類器はターゲットフレーズを、ポジティブ、ネガティブ、ニュートラルに分類する。

5 データセット

BERT-Transformer モデルではアスペクトベースの評判分析のためのマルチラベルデータセットが用いられるが、提案手法にて示したモデルは現在構築中である。そこで、構築が済んでいる部分のモデルの性能を確認するために以下のデータセットを用いて 2 値分類をした。

5.1 評判分析チェック用データ

評判分析チェック用データ²とは、TripAdvisor の日本語レビュー文書の訓練用データと評価用データから作成した評判分析チェック用データである。文章データには、好評（ラベル 1）、不評（ラベル 0）のラベルが付与されている。

6 実験

提案手法にて示したモデルは現在構築中である。そこで、構築が済んでいる部分のモデルの性能を確認するために、本実験では BERT Transformer モデルの一部を用いて 2 値分類をした。図 2 に用いた一部のモデルを示す。

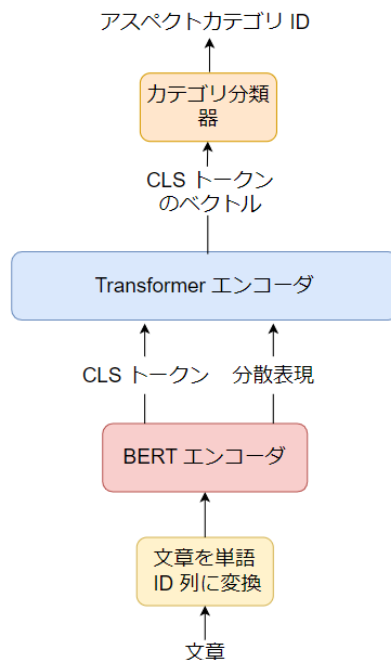


図 2: 実験時の BERT-Transformer の一部のモデル概要図

²<https://github.com/1g-hub/okada/tree/master/tripadvisor>

実験をした目的は関連研究をベースに作成したモデルの一部で 2 値分類をして精度とプログラムの正確さを確認するためである。ベースラインとして BERT に線形層を接続したモデルで 2 値分類をした。また, BERT Layer の最終層と全結合層のみ Fine-Tuning をした。表 1 に実験で用いたデータの内訳を示す。

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
6000	4800	1200

表 2 にシンプルな BERT モデルの実験時のパラメータを示す。

表 2: シンプルな BERT モデルの実験時のパラメータ

パラメータ	値
BERT 層の入力次元数	768
BERT 層の出力次元数	2
バッチサイズ	128
最適化関数	Adam
学習率	0.0001
損失関数	CrossEntropy
エポック数	10

表 3 に BERT-Transformer モデルの実験時のパラメータを示す。

表 3: BERT-Transformer モデルの実験時のパラメータ

パラメータ	値
BERT 層の入力次元数	768
BERT 層の出力次元数	768
Transformer 層の層数	1
Transformer 層の入力次元数	768
Transformer 層の出力次元数	2
バッチサイズ	128
最適化関数	Adam
学習率	0.0001
損失関数	CrossEntropy
エポック数	10

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。表 1 のデータにおける訓練データを 5 分割して交差検証をすることで 5 個のモデルを作成した。図 2 にテスト時の正解率が最も高かったモデルの訓練時の訓練データと検証データの損失の推移を示す。図 3 にテスト時の正解

率が最も高かったモデルの訓練時の訓練データと検証データの正解率の推移を示す。

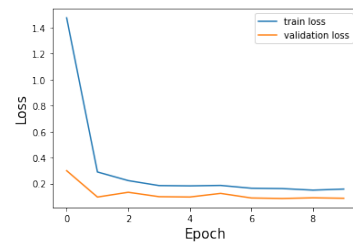


図 3: 実験時の訓練データとバリデーションデータの正解率の推移

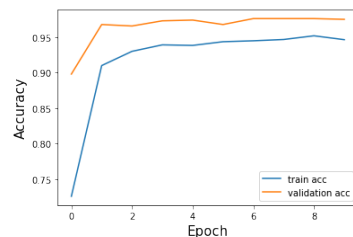


図 4: 実験時の訓練データとバリデーションデータの正解率の推移

図 3 と 図 4 からエポックが進むごとに損失が小さくなっていることや、エポックが進むごとに正解率が上昇していることから正しく学習が進んでいることがわかる。表 4 に最も正解率の高かった訓練モデルにテストデータを入力した時の混合行列を示す。テスト時の正解率は 0.9558 であった。

表 4: BERT-Transformer モデルの最も正解率の高いモデルのテスト時の混合行列

		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	576	25
	ネガティブ	28	571

表 5: ベースラインで最も正解率の高いモデルのテスト時の混合行列

		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	573	28
	ネガティブ	18	581

表 6: BERT-Transformer モデルとベースラインの評価指標の平均と標準偏差

	Accuracy	Precision	Recall	F1
BERT-Transformer	0.9518 (0.0751)	0.9319 (0.1274)	0.9753 (0.09511)	0.9532 (0.0708)
ベースライン	0.9487 (0.0754)	0.9369 (0.1385)	0.9626 (0.1044)	0.9494 (0.0695)

続いて、ベースラインにおいて表 1 の訓練データを 5 分割して交差検証をすることで 5 個のモデルを作成した。表 5 にベースラインのテスト時の混合行列を示す。その時の正解率は 0.9583 であった。表 6 に BERT-Transformer モデルとベースラインのそれぞれのモデルの Accuracy, Precision, Recall, F1 値の平均値と標準偏差を示す。

表 6 から、BERT-Transformer モデルは Precision 以外の評価指標においてベースラインより高い精度で分類できていることがわかる。これらの分類結果が得られたことからプログラムの正確さについての確認はできたといえる。本実験で用いたモデルは構築中であり、Transformer エンコーダ層における Multi-Head Attention 数や、入力文章の Padding 部分の考慮がなされていない。これらの改良をするだけでも更に分類精度が向上することも期待できる。

7 まとめと今後の課題

本研究では新しいモデルとして、BERT-Transformer モデルを提案した。そして、BERT-Transformer モデルの一部の性能を確認するために評判分析チェック用データセットで 2 値分類をした。Precision 以外の評価指標でベースラインを上回ったことから構築中のモデルの分類性能の高さを確認できた。今後の課題はまず、構築中のモデルを完成させることである。また、BERT-Transformer モデルでアスペクトベースの評判分析のためのマルチラベルデータセットを用いる実験を開始する。そして、アスペクトベースの評判分析で日本語文章中の推定結果の理由を抽出するモデルを構築する。

参考文献

- [1] 三浦義栄, 渥美雅保. 事前学習言語モデルを用いたアスペクトベースセンチメント分析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 35 回全国大会 (2021), pp. 2Yin507–2Yin507. 一般社団法人 人工知能学会, 2021.
- [2] 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.
- [3] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Lukasz Kaiser, and Illia Polosukhin. Attention is all you need. In *Advances in neural information processing systems*, pp. 5998–6008, 2017.