

# アスペクトベース評判分析における推定理由フレーズの抽出手法

## 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<sup>1</sup>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。

<sup>1</sup><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 に作成したモデル図を示す.

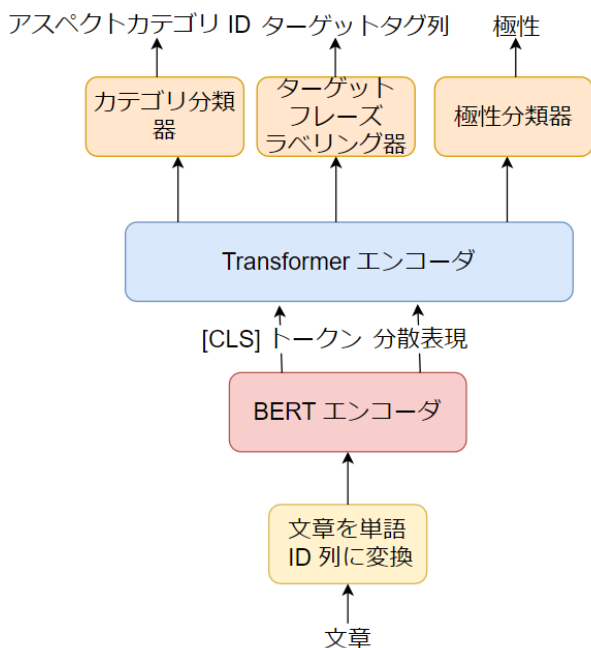


図 1: BERT-Transformer モデル概要図

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

### 5.1 評判分析チェック用データ

さまざまなモデルのチェック用に好評 (ラベル 1), 不評 (ラベル 0) にラベル付けされた 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岡田先生が作成した評判分析チェック用データ<sup>2</sup>を用いた.

## 6 実験

提案モデルの Transformer エンコーダ層の Multi-Head Attention 数は一般的には 8 であるが, プログラムの正確さの確認のため, 今回の実験時では 1 とした. また, Transformer エンコーダ層での Attention の重み計算時には入力文章の Padding 部分の特徴量に関する考慮はしていない. これらの前提のもとで, 実験 1 では BERT Transformer モデルでの 2 値分類をした. 実験 1 をした目的は 2 値分類タスクを通じてモデルの正確さを確認するためである. 実験 2 では BERT 層のみでの 2 値分類をした. 実験 2 をした目的は, 現時点での提案モデルの分類精度の比較対象とすることである. 実験 2 を本実験でのベースラインとして実験 1 と分類精度を比較した. 表 1 に実験 1 と 2 で用いたデータの内訳を示す.

表 1: 2 値分類タスクに用いたデータ

データ	データ数
総データ	6000
訓練データ	3600
バリデーションデータ	1200
テストデータ	1200

### 6.1 実験 1

提案手法にて示した BERT-Transformer モデルは実装したが, ターゲットフレーズラベリング器や極性分類器は, まだ実装できていない. 実験 1 では 2 値分類をしてプログラムの確認やシンプルな分類で

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

どれほど正しく分類できるか確認した. 表 2 に実験時のパラメータを示す.

表 2: 2 値分類タスクに用いたパラメータ

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. 図 2 に訓練時の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. 図 3 に訓練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を示す.

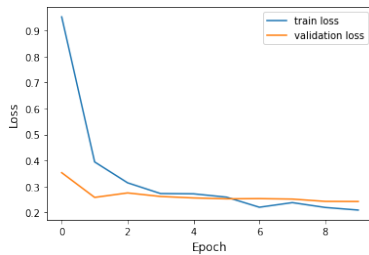


図 2: 実験 1 の訓練データとバリデーションデータの正解率の推移

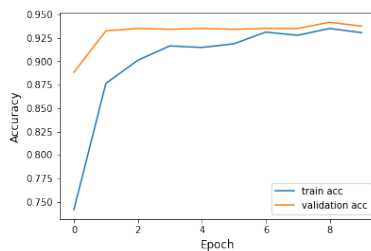


図 3: 実験 1 の訓練データとバリデーションデータの正解率の推移

図 2 からエポックが進むごとに損失が小さくなっていき, 過学習もみられないことから正しく学習が進んでいることがわかる. また, 表 3 からエ

ポックが進むごとに正解率が上昇していることがわかる. 図 4 に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 0.9467 であった.

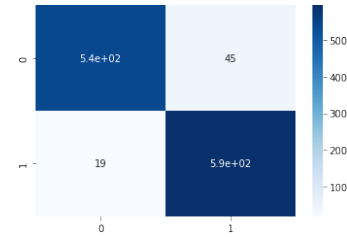


図 4: 実験 1 のテスト時の混合行列

図 4 からテストデータにおいても高い精度で分類できていることが視覚的に判断できる. また, 実験 1 の目的であるプログラムの正確さについての確認ができたと考えられる.

## 6.2 実験 2

実験 2 では BERT モデルに 1 層の線形層を追加したモデルで 2 値分類をした. 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, BERT Layer の最終層と全結合層のみ Fine-Tuning をした. 表 3 に実験時のパラメータを示す.

表 3: 2 値分類タスクに用いたパラメータ

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

図 5 に訓練時の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. 図 6 に訓練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を示す.

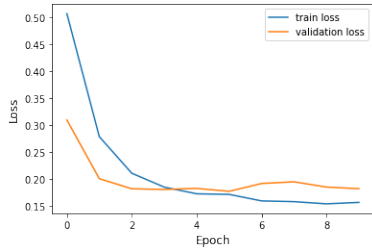


図 5: 実験 2 の訓練時の訓練データとバリデーションデータの損失の推移

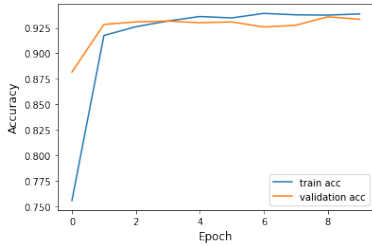


図 6: 実験 2 の訓練データとバリデーションデータの正解率の推移

図 5 からはエポックが進むごとに損失が小さくなっていき、過学習もみられないことから正しく学習が進んでいることがわかる。また、図 6 からはエポックが進むごとに正解率が上昇していることがわかる。図 7 に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  $0.9583$  であった。

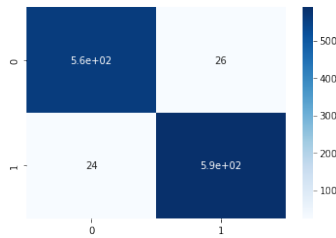


図 7: 実験 2 のテスト時の混合行列

表 4 に実験 1 と実験 2 の正解率を示す。

	実験 1	実験 2
正解率	0.9467	0.9583

実験 2 より実験 1 の正解率が低い原因として、実験時の前提として述べた Transformer エンコーダ層における Multi-Head Attention 数の少なさや、入力

文章の Padding 部分の考慮がないことが考えられる。新モデル提案の目的は分類精度を向上させることではなく、各トークンとアスペクトカテゴリの関連性の導出や Attention の可視化をすることである。これらのことから、分類精度が低下した結果を否定的には捉えていない。しかし、分類精度を更に向上させることは今後モデルを改良する際の重要な点の 1 つであると考えられる。

## 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 モデルを実装した。そして、評判分析チェック用データセットで 2 値分類をした。実験 1 と実験 2 のテスト時の正解率はそれぞれ  $0.9467$ ,  $0.9583$  であった。新しく提案したモデルの正解率の方が低いが、Transformer エンコーダ層における Multi-Head Attention 数を増やして、入力文章の Padding 部分の考慮をすることでこの問題は改善できると考えられる。今後の課題はまず、先述の問題に取り組むことである。また、提案手法にて示した BERT-Transformer モデルでマルチラベルデータセットを用いた実験をする。そして、分類タスクにおける精度の向上や、ターゲットラベリング器と極性分類器の実装などを進めながらモデルの改良を目指す。

## 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.