

敵対的不変表現学習を用いたアスペクトベース感情分析

水谷宏太¹ 南條浩輝²

¹ 滋賀大学大学院データサイエンス研究科 ² 滋賀大学データサイエンス学部

¹ s6022139@st.shiga-u.ac.jp ² hiroaki-nanjo@biwako.shiga-u.ac.jp

概要

文書や文よりも細かいアスペクトを単位とするアスペクトベース感情分析 (ABSA) の研究を行う。既存の ABSA データセットは作成のコストの大きさからサイズは大きくなく、収録されているテキストドメインが限られている。このため、実際の ABSA の分析対象のテキストと同じドメインのデータが得られず、高い精度を得にくいという問題がある。本研究では、ABSA モデルの学習時に敵対的学習を行うことでドメイン不変な表現を獲得し、学習データと異なるドメインでの ABSA の精度の高精度化を目指す。BERT を用いた感情極性分類モデルの学習において、敵対的不変表現学習を導入し、学習時と異なるドメインのデータセットに対する ABSA における敵対的不変表現学習が有効性を示した。

1 はじめに

近年、Instagram や X (旧 Twitter) などの SNS の普及に伴い、テキストデータを対象とする感情分析の重要性が高まっている。より詳細な感情分析のため、文書や文単位よりも粒度の細かい、文中のある対象の一つの側面 (アスペクト) [1] ごとに感情分析をするアスペクトベース感情分析 (Aspect Based Sentiment Analysis: ABSA) が望まれている。

ABSA はその対象の粒度の細かさから、データセット作成が難しい。すなわち、文を対象とした従来の感情分析データセットよりもアノテーションコストが高いというデメリットがある [2]。さらに、ABSA 研究用の一般的なデータセットが収録するテキストドメインは、文レベルの感情分析用のデータセットよりも限定的であり、データセットのサイズも小さい。実際に、日本語における ABSA を目的とした代表的なデータセットに chABSA-dataset¹⁾ があり、収録されている文は 6,119 文である一方、文レベルの感情分析で用いられる代表的なデータセット

である Multilingual Amazon Reviews Corpus [3] の日本語レビュー文が 210,000 件あることを考えるとやはり少ない。

このような背景から、ABSA を行いたいケースにおいて、分析対象のテキストと同じドメインのデータセットがないため、高精度な ABSA が行えないという問題がある。

本研究では、学習データと異なるドメインでの ABSA の精度の高精度化を目指す。具体的には、ABSA モデルの学習時に敵対的学習を行うことでドメイン不変な表現を獲得し、それをを用いた ABSA の研究を行う。

2 アスペクトベース感情分析

ABSA には様々なタスクがある [4]。それぞれのタスクは ABSA に関連する四つのいずれか要素を特定することを目的としている。その四つの要素とは、アスペクト用語 (Aspect Term)、アスペクトカテゴリ (Aspect Category)、意見用語 (Opinion Term)、感情極性 (Sentiment Polarity) である [5]。アスペクト用語とは、与えられたテキスト内において明示的に表れる意見の対象のことである [4]。アスペクトカテゴリは、エンティティとアトリビュートから構成され、エンティティがアスペクト用語が何のカテゴリに属するかを意味し、アトリビュートがその対象のどの属性について言及されているかを示す [6]。なお、アスペクトカテゴリを特定する際は事前に候補となるカテゴリを定義しておく必要がある [4]。意見用語は、テキスト内でアスペクト用語に対する感情を表している表現である [4]。感情極性はアスペクト用語に対する意見用語の感情極性を示し、通常は “positive”, “negative”, “neutral” のいずれかに属する [4]。

1) <https://github.com/chakki-works/chABSA-dataset/tree/master>

3 提案手法

3.1 概要

本研究では、BERT を用いたマルチタスクモデルに対し敵対的不変表現学習を行い、学習データとドメインの異なるデータセットに対する分類精度を検証する。本研究で行う ABSA においては、アスペクト用語とそれを含む文章を事前学習済み BERT に [SEP] で繋げて入力して [CLS] トークンに対応する出力を得る (2 文入力タスク) ことで、文中のアスペクト用語に関する埋め込み表現を抽出し、これを感情極性分類に用いる。事前学習済みモデルとして、日本語のウィキペディアデータで事前学習された東北大 BERT²⁾ を利用した。

3.2 敵対的不変表現学習

敵対的不変表現学習とは、学習に用いるドメイン (ソースドメイン) のデータと分析対象としたいドメイン (ターゲットドメイン) のデータの分布が異なる場合に用いられる敵対的学習手法の一つである。本研究では、敵対的不変表現学習の手法として Ganin et al.(2016) [7] が提案した Domain Adversarial Neural Networks (以下 DANN) を用いる。DANN は、ドメインに依存しない特徴量、すなわちドメインによって不変な特徴量をもとにソースドメインのデータを分類するよう学習することによってターゲットドメインに対する予測精度を下げないようにする学習手法である。

3.3 本研究のフレームワーク

BERT (エンコーダ) でアスペクト単語とそれを含む文の表現 (特徴量) を得たのち、感情極性分類器とドメイン識別器にそれを入力する。感情極性分類器とドメイン識別器についてはそれぞれのタスクが解けるように学習が行われる。BERT については、感情極性分類はできるもののドメイン識別はできなくなるような特徴を取り出すように学習される。ドメイン識別器と BERT が敵対的に働いて学習が行われることで BERT は感情極性分類のためのドメイン不変な特徴量を取り出せるようになる。

2) <https://huggingface.co/cl-tohoku/bert-base-japanese-whole-word-masking>

4 データセット

4.1 chABSA-dataset

chABSA-dataset は 2016 年度の上場企業の有価証券報告書をベースに作成されたデータセットである。各文に対して “positive”, “negative”, “nertral” 感情極性のラベルだけでなく、各ターゲットのアスペクトに対して感情極性のラベルが付与されている。アスペクトカテゴリーは、ターゲットの単語が何のカテゴリに属するかを意味するエンティティ 4 つ (company, business, product, NULL), その対象のどの属性について言及されているかを示すアトリビュート 6 つ (general, sales, profit, amount, price, cost) を組み合わせた 23 種類のアスペクトカテゴリー (company#price が dataset に存在しないため 23 種類) とそれに market#general を合わせた合計 24 個である。各アスペクトカテゴリーはエンティティとアトリビュートを組み合わせて company#general, product#amount といった形になっている。chABSA-dataset では、各文に対してターゲットフレーズとそのアスペクトカテゴリー、極性が付与されている。例えば、「当社グループの主力事業が属するインターネット広告市場は、当年度においても広告市場全体の伸びを上回る成長が続きました」という文に対して、ターゲットフレーズは「インターネット広告市場」、アスペクトカテゴリーは market#general、感情極性が positive というように情報が付与されている。

本研究では、chABSA-dataset 中の感情情報が含まれているアスペクトカテゴリー 7,723 個のうち、ターゲット・文のペアで複数のアスペクトカテゴリー・感情極性が付与されているサンプルを排除した 6,849 個のサンプルを用いる。これは本研究では、ターゲット・文のペアを入力とする際に、アスペクトカテゴリーと感情極性をそれぞれ一つずつ予測するためである。排除したサンプルの例を挙げる。ターゲットが「食品製造販売部門」、文が「食品製造販売部門においては、取引先の見直し等により、売上高は昨年を下回りましたが、原料価格が低下したこと、販売経費の削減により、所定の利益を確保することができました」であるサンプルの場合、文中では「食品製造販売部門」の売上高に関する言及と利益に関する言及がある。このうち、売上高に関しては「昨年を下回りましたが」とあるためネガティブに、利益に関しては「所定の利益を確保することが

表 1 日本語ポストデータセット

ターゲットフレーズ	total	positive	negative	neutral
マイナンバーカード	23	9	7	9
テレワーク	25	9	8	8
ライドシェア	24	9	8	8
インボイス制度	23	7	10	8
ベーシックインカム	25	7	8	10
選択的夫婦別姓	25	9	8	8
total	145	51	46	48

できました」とあるため、ポジティブに述べられていることがわかる。このような場合にデータセットではアスペクトカテゴリが business#sales の場合は negative, business#profit の場合は positive とラベリングしてあるが、今回はターゲットと文のペアに対してアスペクトカテゴリと感情極性をそれぞれ一つずつ予測するため、このようなサンプルは分析対象から排除している。

この chABSA-dataset を学習・検証・テスト用のデータに分割して ABSA における敵対的不変表現学習の効果を調べる。学習データ 5,150 件、検証データ 1,130 件、評価データ 566 となるよう分割した。学習・検証評価データセットごとに異なるアスペクトカテゴリのテキストを割り当てた。これにより、学習に用いていないアスペクトカテゴリのテキストに対する感情極性分類の精度を比較することができる。すなわち、擬似的に未知のドメインのテキストデータに対して感情極性分類を行い、その精度を比較する

4.2 日本語ポストデータセット

次に chABSA-dataset で学習したモデルを用いて、それとは全く別のドメインのデータセットに対して ABSA を行い、敵対的不変表現学習の効果を調べる。本研究では、日本語ポストデータセットを作成した。

6 種類のターゲットフレーズをアスペクト用語として用意し、各ターゲットフレーズについて 25 文ずつ合計 150 文の感情的な意見が述べられている文章を収集した。主著者が “positive”, “negative”, “neutral” がそれぞれおおよそ均等となるよう選択した。さらに各文章に対して別の 2 名のアノテータを割り当て、“positive”, “negative”, “neutral” のラベルを付与してもらい、2 名以上により同ラベルが付与されたデータのみを残した³⁾。

その結果、日本語ポストデータセットの合計数は

3) ただし、positive, positive, negative または positive, negative, negative となった文はデータから除いた

表 2 chABSA-dataset における敵対的不変表現学習の効果

	Accuracy	Macro-F1
GRL+Multi (提案)	0.933	0.834
Multi	0.938	0.814
Single	0.952	0.870

145 となった。データセットの概要を表 1 に示す。

5 実験

本研究では、テキストのアスペクトカテゴリ (エンティティ#アトリビュート) 識別とアスペクトの極性 (positive/negative/neutral) 分類を同時に行うマルチタスクモデルに対して DANN を適用することで異なるドメイン (本研究では学習に使わないアスペクトカテゴリを指す) のデータに対する感情極性の分類精度がどのように変化するかを標準的なマルチタスクモデルや感情極性分類のみを行うモデルと比較する。

5.1 chABSA-dataset での敵対的不変表現学習の効果

感情極性分類のみを行うモデル (以下 Single), カテゴリ識別と感情極性分類を行う通常のマルチタスクモデル (以下 Multi), カテゴリ識別と感情極性分類を行うマルチタスクモデルに勾配反転層を加えて敵対的不変表現学習を行ったモデル (以下 GRL+Multi), を比較した。

エポック数を最大 30 回として各エポックごとに検証データに対する Accuracy を測る。最も Accuracy が高かったエポックのモデルを採用し、評価データに対する精度を求めた。

感情極性分類の精度を表 2 に示す。通常のマルチタスクモデル (Multi) や感情極性分類のみを行ったモデル (Single) でも分類精度が高く、敵対的不変表現学習の効果は見られなかった。この原因として、学習データと評価データがどちらも有価証券報告書を元に作られた同一データセットに含まれるテキストであり、意見を表す表現が似ていたことが考えられる。すなわち、ドメイン不変の特徴量を用いなくても高精度に感情極性分類ができたと考えられる。

5.2 日本語ポストデータセットでの実験

次に学習に使ったデータと異なるドメインのデータである日本語ポストデータセットでの敵対的不変表現学習の効果について述べる。

表3 日本語ポストデータセットにおける敵対的不変表現学習の効果

	Accuracy	Macro-F1
GRL+Multi (提案)	0.462	0.461
Multi	0.448	0.418
Single	0.400	0.400

5.2.1 敵対的不変表現学習の効果

初めに 5.1 節で述べた chABSA-dataset で学習したモデルをそのまま用いて感情極性分類を行い、各モデルの分類精度を比較した。結果を表 3 に示す。

敵対的不変表現学習を行ったモデル (GRL+Multi) が Accuracy, Macro-F1 score とともに最も分類精度が高いという結果になった。5.1 節で行った実験では学習データと評価データが同一データセットに含まれるテキストであり、ドメイン不変の特徴量を用いなくても高精度に感情極性分類ができたが、学習データと意見を表す表現が異なる日本語ポストデータセットに対しては、敵対的不変表現学習により獲得したドメイン不変の特徴量が感情極性分類に有効に働いたことが示唆された。反対に、通常のマルチタスクモデル (Multi) や感情極性分類のみを行ったモデル (Single) に関しては、本実験では敵対的不変表現学習を行ったモデルと比較すると精度が低くなっている。これは、異なるドメインのデータに対しても学習時のデータにおける意見を表す表現に依存してしまい、上手く分類が行えなかったことが考えられる。学習データと異なるドメインのデータに対しての感情極性分類において敵対的不変表現学習の有効性が示唆された。

5.2.2 少量ドメインデータでのドメイン適応の効果

次に 5.1 節で述べた chABSA-dataset で学習したモデルに対して、少量のドメインデータで fine-tuning を行ってドメイン適応した。具体的には日本語ポストデータセットの一部を使って追加学習を行い、日本語ポストデータセットに対して感情極性分類を行った。日本語ポストデータセットをターゲットフレーズごとに 6 分割し、学習データ：検証データ：評価データに 4:1:1 の比率で割り当てた。全てのデータが一度は評価データとして利用されるように割り当てのプロセスを 6 回繰り返し、各評価データに対する精度の平均値を取るクロスバリデーションを実施した。

エポック数を最大 10 回とし、クロスバリデーションの 1 回目の検証データに対して Accuracy が最大となるエポック数のモデルを採用した。残りの 5

表4 少量データによる追加学習における敵対的不変表現学習の効果 (日本語ポストデータセットで評価)

	Accuracy	Macro-F1
GRL+Multi (提案)	0.625	0.615
Multi	0.569	0.529
Single	0.597	0.576

回のクロスバリデーション時は最初に決めたエポック数に固定して学習を行った。結果を表 4 に示す。

追加学習により、いずれのモデルも 5.2.1 節で示した結果よりも精度は向上した。ここでも 5.2.1 節の結果と同様に敵対的不変表現学習を行ったモデル (GRL+Multi) がいずれの指標においても最も分類精度が高かった。

対象としたいドメインにおける少量の学習データが利用できる場合でも、あらかじめ別ドメインで敵対的不変表現学習をしたモデルを利用することの有効性を確認した。

6 おわりに

ABSA の研究を行った。分析対象としたいドメイン (ターゲットドメイン) のラベル付きデータが不足しているという課題に対して、既存のデータセットの学習データとしての活用法、具体的には敵対的不変表現学習の有効性を検討した。ターゲットドメイン用の学習データがない場合、および少量の場合において、敵対的不変表現学習の有効性を示した。

本研究では、一つのアスペクト用語と文章のペアに対して複数のアスペクトカテゴリと感情極性が割り当てられているデータは分析対象から除いたが、このデータも扱うことは今後の課題である。今回用いた DANN 以外の敵対的不変表現学習の調査も今後の課題である。

さらに、BERT をはじめとする自然言語理解モデル自体の持つ課題への対応も今後の課題である。文章による意見の表現時に、「賛成だ」「反対だ」と言ったような直接的な表現ではなく、修辞法を用いられることがある。例えば、「賛成すると思う?」、「賛成するわけなくない?」、「なぜ賛成できるのかわからない」など、negative な意見を主張する際に positive な単語と疑問や否定表現を組み合わせるといったことが行われる。これらは negative な意見を述べているにも関わらず、表層的には「賛成」という positive な語がつかわれており、自然言語理解モデルが苦手とする問題といえる。

参考文献

- [1] 張懿陽, ラファウ・ジェブカ, 荒木健治. BERT モデルと補助文自動生成に基づいた日本語アスペクトベース感情分析の精度向上. ことば工学研究会: 人工知能学会第 2 種研究会ことば工学研究会資料, Vol. 70, pp. 83–89, 12 2022.
- [2] Yiming Zhang, Min Zhang, Sai Wu, and Junbo Zhao. Towards Unifying the Label Space for Aspect- and Sentence-based Sentiment Analysis. In Smaranda Muresan, Preslav Nakov, and Aline Villavicencio, editors, **Findings of the Association for Computational Linguistics: ACL 2022**, pp. 20–30, Dublin, Ireland, May 2022. Association for Computational Linguistics.
- [3] Phillip Keung, Yichao Lu, György Szarvas, and Noah A. Smith. The Multilingual Amazon Reviews Corpus. In **Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing**, 2020.
- [4] Wenxuan Zhang, Xin Li, Yang Deng, Lidong Bing, and Wai Lam. A Survey on Aspect-Based Sentiment Analysis: Tasks, Methods, and Challenges. **IEEE Transactions on Knowledge and Data Engineering**, 2022.
- [5] Wenxuan Zhang, Yang Deng, Xin Li, Yifei Yuan, Lidong Bing, and Wai Lam. Aspect Sentiment Quad Prediction as Paraphrase Generation. In Marie-Francine Moens, Xuanjing Huang, Lucia Specia, and Scott Wen-tau Yih, editors, **Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing**, pp. 9209–9219, Online and Punta Cana, Dominican Republic, November 2021. Association for Computational Linguistics.
- [6] 三浦義栄, 赤井龍一, 渥美雅保. 文中の複数アスペクトのセンチメント分析のための自己注意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 34 回 (2020), 3Rin441, pp. 1–4. 一般社団法人人工知能学会, 2020.
- [7] Yaroslav Ganin, Evgeniya Ustinova, Hana Ajakan, Pascal Germain, Hugo Larochelle, François Laviolette, Mario March, and Victor Lempitsky. Domain-Adversarial Training of Neural Networks. **Journal of machine learning research**, Vol. 17, No. 59, pp. 1–35, 2016.