

前後文のアスペクトを考慮した レビュー文に対する教師ありアスペクト推定の高精度化

薦永 紗希[†]木村 優介[‡]寺本 優香^{††}波多野 賢治[†][†] 同志社大学文化情報学部[‡] 同志社大学大学院文化情報学研究科^{††} 同志社大学文化遺産情報科学調査研究センター

1 はじめに

レビューは消費者の意思決定において有用な情報源である。しかし、それらすべてに目を通すことは困難であるため、商品や店舗の評価項目であるアスペクト¹ごとに要約し、それらを提示する必要がある。そのため要約の作成には、レビュー文からアスペクトを推定する技術が求められる。

既存手法の多くは特定のアスペクトと関連のある語の発見に依存している。しかし、そのような語を含まないレビュー文に対しては既存手法が有効とは言えず、語以外に新たなアスペクト推定の手掛かりが必要である。

語以外のアスペクト推定の手掛かりとして、それらの関係がある。特定のアスペクトと関連のある語を含まない文に出現するアスペクトに着目した研究では、あるアスペクトを含む文の周辺文は、同じアスペクトで書かれる場合があるという報告がある [1]。このことから本研究では、周辺文に出現するアスペクトがアスペクト推定の手掛かりとなると考え、周辺文に出現するアスペクトを考慮することで、従来よりも正確にアスペクト推定が可能となる手法を提案する。

2 先行研究

既存手法では、事前学習済み言語モデルである BERT を用いて、アスペクトとレビュー文のペアに関連があるかどうかを分類することによりアスペクトを推定している [2]。BERT は事前学習で大量の語彙を学習していることから、高い精度でアスペクト推定が行えたと報告されている。しかし、この手法は推定対象のレビュー文に含まれる語しか考慮していないため、アスペクトに関連する語が該当文中に出現していない場合の推定が難しい。

3 提案手法

我々の基礎分析によって、アスペクトにより前後いずれかの文に出現しやすいアスペクトが異なることがわかった。そのため、語の分散表現から計算された文のベクトルと各アスペクトの出現確率を表現するベクトルの両方を使用することで、一文のアスペクト推定の精度を向上させる手法を提案する。

アスペクトを表すベクトルは、図 1 のように、先行研究の BERT のペア分類モデルから得る。ペア分類モデルは、入力レビュー文中に、入力されたアスペクトが出現していれば「関連あり」、していなければ「関連なし」というラベルを出力するように学習しており、出力層の Softmax 関数から各ラベルに分類される確率を得ることができる。「関連あり」というラベルである確率を各ア

¹ レストランのレビューであれば、食事や立地などがアスペクトに該当する。

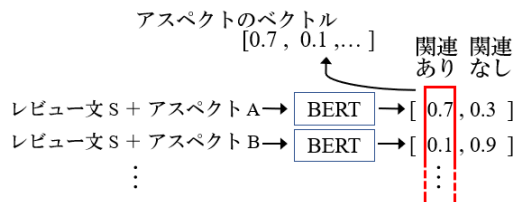


図 1 アスペクトのベクトルを取得する方法

スペクトがレビュー文中に出現している確率として取得し、各アスペクトの出現確率を要素とするベクトルを、アスペクトのベクトルとして使用する。

また、アスペクトだけでなくレビュー文に含まれる語の情報も考慮するため、BERT から得た各語のベクトルを平均化したベクトルを、文のベクトルとして使用する。こうして得たアスペクトのベクトルと文のベクトルを連結したものを、レビューの中に含まれている一文のベクトルとし、その前後文から得たベクトルも連結することで、既存手法では考慮できていなかった前後文を考慮するように設計し、アスペクト推定時の特徴量として用いる。

4 評価実験

前後文に出現するアスペクトを考慮することの有効性を確認するため、前後文を考慮していない先行研究と提案手法のモデルにデータセットを用いて学習させ、テストデータを用いて推論した際にアスペクトを正しく推定できているかどうかを比較する。評価指標には、先行研究と同様に F_1 のマイクロ平均を用いる。 F_1 とは適合率と再現率の調和平均であり、文から多くのアスペクトを推定しかつその正確さを表しているが、そのマイクロ平均を用いることで、極端にデータ数が少ないアスペクトの影響を小さくして評価することができる。

5 おわりに

本研究では、レビュー文のアスペクトを推定する際にその前後文のアスペクトをも考慮する方法を提案した。予備実験ではアスペクトのベクトルの学習に使用したレビュー文数が 2,500 と少なかったため、提案手法のマイクロ F_1 値が先行研究よりもわずかに低いという結果が得られたが、その数を増やすことで精度の改善が図られるかどうかを確認する予定である。

参考文献

- [1] H.-Y. Chen and H.-H. Chen. Implicit polarity and implicit aspect recognition in opinion mining. In *Proc. of ACL2016*, Vol. 2, pp. 20–25. ACL, 2016.
- [2] M. Hoang, et al. Aspect-based sentiment analysis using BERT. In *Proc. of NoDaLiDa 2019*, pp. 187–196. Linköping University Electronic Press, 2019.