

旅行サイトにおけるレビュー文のカテゴリを考慮した評価値予測

井上 敬通†

藤田 桂英‡

†東京農工大学 工学部

‡東京農工大学大学院 工学研究院

1 はじめに

近年、ユーザがインターネット上でレビューを投稿することは、身近なものとなっている。ユーザレビューの数が増加することで、より高度なユーザ分析が可能になるが、膨大なレビューを手で正確に分析することは難しいため、高速で精度の高いレビュー文の解析が必要である。

本研究では、旅行サイトにおいて、レビュー文から複数の評価値をより高精度に予測することを目的として、レビュー文から評価値を予測する手法を提案する。さらに、評価実験により、提案手法が既存手法と比較して予測精度が上回ることを示す。

2 楽天トラベルデータにおける評価値予測

2.1 楽天トラベルデータセット

本研究では、楽天トラベルデータ [1] を用いる。楽天トラベルのユーザは6つの評価カテゴリをそれぞれ5段階で評価し、レビュー文とともに投稿する。ここで、6つのカテゴリとは「サービス」「立地」「部屋」「設備・アメニティ」「風呂」「食事」である。楽天トラベルのレビュー文を用いた評価値予測では、まず、レビュー文の分散表現を取得し、その分散表現を入力とする分類器を用いて、多クラス分類問題を解く。

2.2 PV-DM と全結合層による手法

外山ら [2] は、1レビューに対し、文全体と1文毎の分散表現をそれぞれ生成し、複数カテゴリの評価点を同時に予測する手法を提案している。分散表現を取得するために、Distributed Memory model of Paragraph Vector (PV-DM) を利用することで、レビュー文中の単語の位置関係を考慮して、評価値を予測する。さらに、文毎の分散表現を取得し、文の相対位置によって重み付けした加重平均を用いることで、文の位置関係を考慮した、評価値の予測には、分散表現を入力とするニューラルネットワークを用いた。評価実験の結果から、単語及び文の位置関係を考慮することが有効であると報告された。しかし、この手法では、分散表現

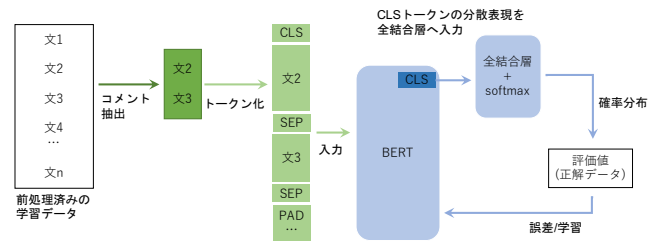


図 1: 提案手法の概観

とクラス分類の学習がそれぞれ独立しているため、学習の柔軟性に課題がある。

2.3 カテゴリ毎の文抽出による評価値予測

張 [3] は、レビュー文から評価カテゴリに言及している部分を抽出して予測を行なった。レビュー文の前処理をしたあと、「評価カテゴリ名の類義語」や「類義語の並列表現」などのキーワードリストを作成し、コメントを抽出した。その後、「自立語」「自立語の否定」「評価語」「評価語の否定」という4つの素性と、それぞれの重みを定義し、素性を要素にもつ重み付きの Bag-of-Words を分散表現として取得し、L2 正規化ロジスティック回帰を用いて、評価値を予測した。張は、評価カテゴリについて言及していないレビューを除外することが、モデルの性能向上に有効であると、述べている。また、「設備・アメニティ」についてはキーワードリストが不十分であるため、高い精度が出なかったことも報告している。課題として、分散表現をルールベースで作成していること、分類器が古典的であることが挙げられる。

3 旅行サイトのレビュー文における評価値予測のための BERT モデル

本研究では、分類器として BERT を用いた評価値の予測を行う手法を提案する。BERT とは、Google が 2018 年に公開した、自然言語処理における機械学習モデルである [4]。提案手法の概観を図 1 に示す。BERT の入力は、文字列をトークン化した系列であり、入力系列の長さは固定である必要がある。そのため、入力トークンの数が不足していた場合、不足分を [PAD] で埋める。一方で、超過していた場合、入力文の省略を行うことで数を揃える。先頭に [CLS]、文末に [SEP] とい

Rating Prediction Considering Categories for Travel Review

†Faculty of Engineering, Tokyo University of Agriculture and Technology

‡Institute of Engineering, Tokyo University of Agriculture and Technology

うトークンを挿入し、[CLS]の出力を入力文の分散表現として学習する。

BERTは、Transformer[5]のEncoderを連結した構造となっており、距離の離れたトークンどうしの関連度を考慮して学習することが可能である。Transformerの特徴であるAttention層は、入力トークンどうしの関連度を距離に関係なく取得するための構造である。以下に、Attentionの一般式を示す(式1)

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{Q \cdot K^T}{\sqrt{d_k}}\right) \cdot V \quad (1)$$

ここで、Q (Query), K (Key), V (Value)は $d_k \times \{\text{シーケンス長}\}$ の行列である。Q・K^TはQに対するK毎の関連度を表し、 $\sqrt{d_k}$ で割ることで、スケールリングする。 $\frac{Q \cdot K^T}{\sqrt{d_k}}$ を、Qに対するK毎の関連度(=重み)とし、softmax関数で正規化したのち、Vの重み係数とする。この重み係数により、トークン毎に「注目するトークン」を判別して学習を行うことができる。

BERTはテキストデータで事前学習した後、ドメインに合わせた訓練データで転移学習を行って使用するが、提案手法では、日本語で事前学習されたBERT[6]を用いる。訓練データは、張の手法を参考にし「予測対象のカテゴリに言及しているレビュー全文」を抽出する。学習では、評価カテゴリ毎に分類器を用意し、それぞれ転移学習を行う。

4 評価実験

外山らの手法をベースライン1、張の手法をベースライン2として、提案手法と比較する。実験データは、楽天トラベルデータのうち、最新の420,766件のレビューを使う。カテゴリ毎に抽出したデータの件数は、「サービス: 194,086件」「立地: 128,866件」「部屋: 237,463件」「設備・アメニティ: 121,832件」「風呂: 195,404件」「食事: 292,735件」である。

提案手法のパラメータは、Attention層と隠れ層のドロップアウト率が0.1、入力層の次元が512次元、中間層の次元数が768次元、Attention層の数が12層、損失関数が交差エントロピー誤差、学習率が 5.0×10^{-5} とした。また、学習では、バッチサイズを290、入力系列の長さを100としている。

表1に実験結果を示す。BL1はベースライン1、BL2はベースライン2である。ベースライン1は評価カテゴリ毎に予測を行っていないため、平均の精度のみ記載している。提案手法は、全てのカテゴリについてベースライン2を上回り、全体の精度は2つのベースラインを上回った。提案手法がベースライン1の精度を上回ったことから、分散表現とクラス分類をEnd-to-End

表1: 分類器の精度比較

	BL1	BL2	提案手法
サービス	-	52.48	59.90
立地	-	54.42	58.50
部屋	-	47.42	54.90
設備	-	39.87	47.44
風呂	-	49.71	54.85
食事	-	52.53	59.24
平均	36.86	49.41	55.81

に学習するBERTが、評価値の予測に有効であることが示された。さらに、ベースライン2の精度を上回ったことから、分散表現をルールベースで作成し、ロジスティック回帰で評価値を予測するよりも、BERTでの予測がより高精度であることがわかった。

5 まとめ

本論文では、BERTを用いて、旅行サイトにおけるレビュー文から評価値の予測を行なった。評価実験の結果、BERTによる予測が既存手法よりも高い精度であることを示した。

謝辞

本研究では、国立情報学研究所のIDRデータセット提供サービスから、楽天株式会社による「楽天データセット」を利用させていただきました。

参考文献

- [1] 楽天株式会社:楽天トラベルデータ. 国立情報学研究所情報学研究データリポジトリ.(データセット), 2016. <https://doi.org/10.32130/idr.2.2>
- [2] 外山洋太, 三輪誠, 佐々木裕. 文書・文間及びカテゴリ間の関係を考慮したレーティング予測. 言語処理学会 第22回年次大会, pp.158-161, 2016.
- [3] 張博. レビューテキストの書き手の評価視点に対する評価点の推定. 北陸先端科学技術大学院大学情報科学研究科 修士論文, 2017. <https://dspace.jaist.ac.jp/dspace/handle/10119/14154>
- [4] Jacob Devlin, Ming-Wei Chang, Kenton Lee, Kristina Toutanova. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. In North American Association for Computational Linguistics, 2019.
- [5] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Lukasz Kaiser, Illia Polosukhin. Attention is all you need, in Advances in Neural Information Processing Systems, pp.6000-6010, 2017.
- [6] 柴田 知秀, 河原 大輔, 黒橋 禎夫. BERTによる日本語構文解析の精度向上. 言語処理学会 第25回年次大会, pp.205-208, 2019.