

P07 否定的レビューに対して釈明を含む返信の分類精度の向上

半谷 太一（茨城大学大学院理工学研究科情報工学専攻）

堀田 大貴（茨城大学）

1.背景

- ホテル宿泊予約サイトにはレビュー機能とレビューに対しての返信機能が備えられている場合が多い
- 否定的な表現を含むレビューは、潜在顧客である読者の対象ホテルに対するイメージを低下させる一因である
- 否定的なレビューに対する返信の中には、謝罪をしているだけのものや読者の印象が良い方向に転化する内容のものがある
- 本論文では、後者を「釈明を含む返信」と呼ぶ
- しかし、全ての返信が閲覧されるわけではないため、釈明を含む返信をしていても読者が確認しない場合がある
→釈明をしている場合には、読者に分かるように付記する必要がある

2.目的

現状：大量のデータセットから見つけ出すのは困難
目的：「釈明を含む返信」を自動で推定する

「釈明」は4種類定義（中村ら（2018））

- 誤解の解消 「駐車場は〇〇にございます」
- 解決策の提示 「加湿器が必要な場合はフロントまでお声掛けください」
- 改善の完了 「エアコンの修理が完了いたしました」
- 対応できない理由の説明 「集中暖房のため、お部屋での温度調整ができません」

3.提案手法

Linear-chain CRFとBERTを用いて分類器を構築する

分類器は主に4段階で構築

- レビュー文書と返信文書を「。」「?」「!」で複数の文に分割
- Linear-chain CRFモデルによる返信文のラベル付け
- BERTモデルによるレビュー文のラベル付け（「ポジティブ」「ネガティブ」「中立」）
- ②と③の結果を用いて「釈明を含む返信」をしているデータを推定する

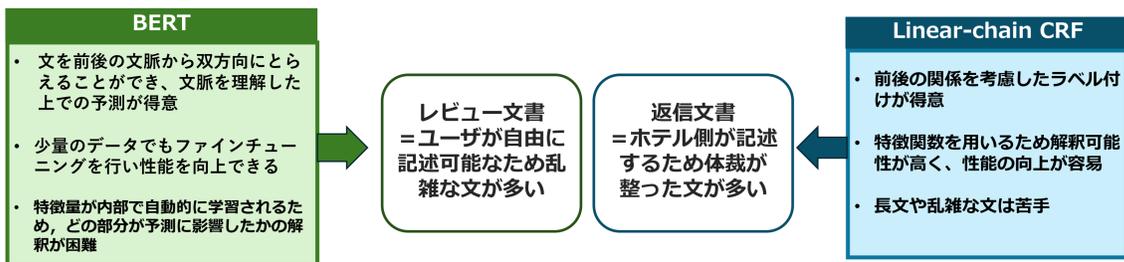


表1：返信文のラベル

ラベル区分	ラベル
<宿泊御礼>	THX-STAY
<投稿御礼>	THX-POST
<提案・指摘感謝>	THX-PROP
<ほめや感想への感謝や提案・指摘感謝>	THX-THOUPROP
<結びでの御礼>	THX-LAST
<ほめへの感謝>	THX-COM
<心がけ・決意>	TRY
<再泊願い>	REC-REP
<署名・フッター>	NAME
<末尾の感謝と署名>	THX-LASTNAME
<定期的挨拶>	STA-GREET
<レビューに対する満足>	SATIS
<お詫び>,<恥じ入り>	SRY
<誤解の解消>	CLEAN (誤解の解消)
<事実述べ>	STAT-FACT
<具体的対応明示>	CORR-SPE (改善の完了)
<抽象的対応明示>	CORR-ABS
<具体的検討明示>	CON-SPEC
<抽象的検討明示>	CON-ABS
<対処明示>	HOWTO (解決策の提示)
<了承願い>	PLZ-UND (対応できない理由の説明)
<対話>	TALK
<情報追加>	ADDINFO
<その他>	O

図1：提案手法概要

4.実験

■データセット

- 2019年の楽天トラベルデータからレビューと返信の数が多い上位3件の宿泊施設を選択し、それぞれのホテルから各500件ずつランダムで選んだ。また、テストデータの種類を増やすため、テストデータに100件のデータを追加（データセットは計1600件）
- それぞれのホテルのデータ500件のうち、各データは300件を訓練データ、100件を検証データ、100件をテストデータとして使用

■モデル

- Linear-chain CRFモデル：表3のように特徴関数を定義して学習させた
- BERTモデル：Hugging Face上にあるAmazonのレビュー20000文で学習された感情分析用のモデルをファインチューニングした

■評価メトリクス

- 適合率（精度）、再現率、F値を用いた

表3：特徴関数

特徴関数
当該文の文種類
文のポジティブスコア
文のニュートラルスコア
文のネガティブスコア
1個前の文種類
1個後の文種類
文の長さ
動詞の有無
過去形の動詞の有無

表2：使用データの内訳

データの種類	否定表現を含むレビューの数	釈明を含む返信の数
訓練データ	367件	96件
検証データ	115件	25件
テストデータ	198件	83件

5.実験結果

以下の3つの結果について評価メトリクスを計算した

- Linear-chain CRFによる返信文書の釈明の有無の推定結果
- BERTによるレビュー文書の否定表現の有無の推定結果
- Linear-chain CRFとBERTによる返信文書の釈明の有無の推定結果

	①の結果	②の結果	③の結果
適合率	0.96	0.91	0.96
再現率	0.65	0.95	0.61
F値	0.78	0.93	0.75

■実験結果

- Linear-chain CRFだけを使用した①の方が③よりも結果が良好である
- F値が0.75という高い値を示した。これは、本論文と同様のテーマの既存研究に比べて0.3~0.5高かった（既存研究ではF値が0.2~0.4）

6.考察・今後

- CRFによる分類では、ラベルの推定が誤っているデータが多かった

- 表4より「改善の完了」が誤りやすいと分かる
考察：「CON-SPEC」や「CORR-ABS」は「改善の完了」と似たような文が多く、文中での出現順序も似ているためであると考えられる

- BERTの結果は値が異常に高いため過学習を起こしている可能性がある

- 今後の展望：実験に用いたデータのバリエーションが少ないため、データの量や種類を増やすことが必要であると考えられる

表4：誤ったラベル推定

正しい文種類	推定した文種類	数
具体的対応明示	具体的検討明示	11
具体的対応明示	抽象的対応明示	8
具体的対応明示	事実述べ	1
具体的対応明示	宿泊御礼	1
誤解の解消	レビューに対しての満足	4
誤解の解消	抽象的対応明示	2
誤解の解消	具体的検討明示	1
誤解の解消	宿泊御礼	1
対処明示	心がけ・決意	1