

評価視点別の言及度を用いた意見文の分類手法の提案

新井智也

近年、ユーザの商品購入や企業のマーケティングに役立つ Web 上の評判情報が増え続けている。評判情報は、対象となる商品全般についてではなく、商品の側面となる様々な評価視点から書かれていることが多い。一般に評価視点は「デザイン」や「携帯性」などの語で表されている。評判情報を閲覧するユーザは、特定の評価視点からその製品の評判情報を知りたい場合でも、すべての評判情報の中からその評価視点に対して言及している文を探さなければならない。このため、ユーザは、増え続ける評判情報をすべて閲覧しなければならず、特定の評価視点について知りたい場合でも、多くの時間を要し効率が悪い。

本研究は、評価視点に基づいて意見文を分類することで、ユーザの評判情報の閲覧の支援を目的とする。評価視点の語を直接含む文を抽出する方法では十分な再現率が得られないことから、本研究では、意見文が評価視点に対して言及している詳しさを言及度として定義し、各評価視点に対する文の言及度を用いて分類を試みる。

本研究で提案する分類手法では、評価視点ごとに関連語辞書を作成し、関連語辞書を用いて算出した言及度から意見文を分類する。関連語辞書は、評価視点の語と共起する語を関連語とし、関連語が評価視点と共起する確率を使って、関連語ごとに関連度を登録しておく。評価視点に対する言及度は、各評価視点の関連語辞書を使って、文に現れる関連語の関連度の総和とする。また、分類された文をユーザに提示するときには、言及度の高い文からソートして提示する。これは言及度が高い文ほど、その評価視点について詳しく述べられていることが多く、ユーザにとってより役立つ文であると考えたためである。

実運用されているレビューサイトの評判情報を用いて、提案法を実装評価した。従来法の SVM を使った意見文の分類法と比較し、本手法の有効性を確認した。本手法による分類の精度は約 75% であり、SVM 法と比較して 7% 程度精度が向上することが分かった。提案法により、「ヘアライン加工の天板がいい感じです。」が「デザイン」という評価視点に、「やはり 700g 代は軽いです。」が「携帯性」という評価視点に分類できるなど、評価視点の語を含まない文も正しく分類されている。

実装した分類手法では、一つの文は一評価視点に分類するとしているが、「デザインが良く、バッテリーも長持ちです。」のような複数の評価視点に言及している文の扱いについても検討していく必要があると考えられる。

(指導教員 佐藤 哲司)