

Title	他者のコメントの引用を考慮したオピニオンマイニング
Author(s)	岡山, 有希
Citation	
Issue Date	2013-03
Type	Thesis or Dissertation
Text version	author
URL	http://hdl.handle.net/10119/11340
Rights	
Description	Supervisor:白井清昭, 情報科学研究科, 修士

他者のコメントの引用を考慮したオピニオンマイニング

岡山 有希 (110015)

北陸先端科学技術大学院大学 情報科学研究科

2013 年 2 月 06 日

キーワード: オピニオンマイニング, 極性判定, 引用文除去, 評判分析.

近年では, ブログやマイクロブログといった個人が情報を発信するメディアが一般的となり, ウェブ上から大規模テキストが容易に得られるようになった. ウェブテキストの中でも, 特にブログ記事は, さまざまなトピックについて個人の思想や感情, 評判などを表現しているため, 世間の人々がトピックに対してどのような考えを抱いているのか知ることができる貴重な情報源として利用できる. しかし, ブログ記事ではしばしば他者のコメントが引用される. 引用された他者のコメントは, ブログ著者がトピックに対してどのような立場をとっているのかを自動的に判断する際に誤判定を引き起こす要因となりうる. そこで, 本研究では, あるトピックに対して記述されたブログ記事をウェブから取得し, ブログ記事中から他者の記事やコメントを引用した箇所を検出・除去した後に, ブログの内容がトピックに対して賛成的立場・反対の立場・中立的立場のどれに当てはまるのかを分類し, 賛成・反対の立場をとるブログ件数を集計するとともに, 賛成意見文, 反対意見文をユーザーに提示することで, 世間の意見を俯瞰的に見ることができるシステムを構築することを目指す. オピニオンマイニングに関する過去の研究では, 他者の記事やコメントの引用に対して特別な処理は行われていない. これに対し, 本研究では, トピックに対する賛成・反対意見を集約するオピニオンマイニングシステムにおいて, 他者のコメントの引用を適切に処理する手法を確立することを目的とする. 本研究で提案するシステムは「コメント収集タスク」「引用箇所判定タスク」「極性判定タスク」の3つのタスクから構成されている. コメント収集タスクでは, 調査対象とするトピックのブログテキストをウェブから取得する. 引用箇所判定タスクでは, 他のウェブサイトから取得し転載されたテキストを検索し除去する. 極性判定タスクでは, ブログの著者がトピックに対して賛成(肯定的)か反対(否定的)か中立のいずれの立場を取るかを判定する. 提案システムの評価実験の結果, 他者の記事やコメントの引用箇所を検出, 除去することにより, ブログ記事の極性判定の精度が3%向上することが確認できた. 個人の意見の集約を図るオピニオンマイニングシステムでは, いかにして著者の意見文を取得するのかが重要である. ブログ記事のように著者の記事と他人の記事やコメントが混在することの多いテキストにおいて, 引用箇所を推定し除去することがオピニオンマイニングの際に有用であるといえる.