

システム情報工学研究科修士論文概要

年 度	平成 27 年度	学位名		修士(工学)
専 攻	知能機能システム	専攻	著者氏名	佐藤 裕也
指導教員氏名 掛谷 英紀				
論文題目 Web 上の言語資源を用いた先見性のある人物の特徴分析				
論文概要 <p>本研究では先見性のある人物と先見性のない人物を区別する指標として Amazon のカスタマーレビューを使用した機械学習を提案した。Amazon において発売当初と比較してその評価が大きく変動した商品にレビューをしているユーザを、先見性のあるカテゴリと先見性のないカテゴリに分類して分析をした。その結果、レビュー対象商品のカテゴリには大きな差が見られた。先見性のあるユーザは先見性のないユーザに比べ、本（洋書を含む）をより多く読む傾向が見られ、中でも文学・評論作品や漫画・ライトノベルを多く読んでいる。機械学習はカテゴリの差による影響を防ぐため本に限定して行い、判定するデータセット一件あたりの文章量を増やすため一つのデータセットをレビュー10件、20件、30件でまとめた場合のそれぞれで行った。ただし、同一ユーザのレビューの使用は130件までとしている。クロスバリデーションの分割数は3,5,7,10,12,15,18のそれぞれで実験をした。一つのデータセットを30件にし、18分割クロスバリデーションにより最大77%の正解率が得られた。1つのデータセットの件数とクロスバリデーションの分割数を増やすことによる正解率の上昇が示されたので、今後標本データを増やすことでより高い正解率を得られると考えられる。機械学習で得たα値の高い素性から、各カテゴリに関して特徴的な素性を得た。先見性のあるユーザのレビューには本の感想や他ユーザへの推薦に言及したものが多く見られた。一方、先見性のないユーザのレビューには本を手にした経緯に言及したものが見られ、政治問題や原子力問題に関心のある人物が多く見られた。今回は評価が大きく変化した本として11冊に絞ってレビューを収集したが、評価が大きく変化した本を自動的に判定・収集することができれば標本データが増えて正解率を高くできる可能性がある。先見性のある人物と先見性のない人物の特徴が明らかになることで、先見性のある人物への投票や先見性のある経営トップのいる会社への就職などの支援ができると考えられる。</p>				
審査日 平成 28 年 1 月 28 日				
審査員	(大学名 職名)	(学位)	(氏名)	
主査	筑波大学 准教授	博士(工学)	掛谷 英紀	
副査	筑波大学 教授	博士(工学)	宇津呂 武仁	
副査	筑波大学 教授	工学博士	森田 昌彦	