

氏名	角田 孝昭		
学位の種類	博士（工学）		
学位記番号	博 甲 第 7701 号		
学位授与年月日	平成 28年 3月 25日		
学位授与の要件	学位規則第4条第1項該当		
審査研究科	システム情報工学研究科		
学位論文題目	顧客レビューテキスト解析に基づく文書作成支援に関する研究		
主査	筑波大学 教授	博士（工学）	山本幹雄
副査	筑波大学 教授	博士（工学）	狩野 均
副査	筑波大学 教授	博士（工学）	福井和広
副査	筑波大学 准教授	博士（工学）	乾 孝司
副査	筑波大学 准教授	博士（情報学）	関 洋平
	(図書館情報メディア研究科)		

論文の要旨

本論文は、文書作成支援を行うシステムの基本手法に関する研究をまとめた論文である。特に、電子商取引 web サイトにおいて商品販売者が消費者向けの文書を作成する過程で、「何を書くべきか」の内容を提案する手法の研究を行っている。大きく2種類の支援タスクを設定し、それぞれのタスクに対する基盤手法を研究した。

1つ目は、消費者に対して商品紹介文を作成するときに商品の紹介に相応しい商品特長を顧客レビューデータから自動で抽出し、書き手に提示するタスクである。販売者は消費者にとって最も魅力的な特長を過不足なく記載することが求められるが、消費者が真に求める特長を見定めることは簡単ではない。本論文では、対象商品と他商品集合の顧客レビューを比較・分析した結果に基づき、商品紹介文に相応しい商品特長を自動抽出する手法を議論した。

2つ目のタスクは、消費者からのメールやレビューに対して販売者が返答を書くときの支援タスクである。基本技術として、対となる二つの文書間における文レベルでの対応関係を自動推定する手法を提案したり、これによって販売者の書いた返答に回答漏れがないか等のアドバイスが可能となる。

評価実験では、実際の電子商取引 web サイトに掲載されている大規模なレビューテキストを用いて評価を行っている。1つ目の手法については、自動抽出された商品の特長が前もって与えられている正解としての特長をどのくらいの精度でカバーしているかを評価し、ナイーブな手法よりも統計的に有意に優れていることを明らかにした。2つ目の手法については、1000文書に対して人手で決めた文対応をどのくらいの精度で正しく認識できるかを評価し、ナイーブな手法よりも提案手法が有意に優れていることを明らかにした。

審査の要旨

【批評】

これまでの文書作成支援の研究では、スペルミス等の表層的な校正を支援するものが多かったが、本研究では文書の内容にまで踏み込んだ作成支援を行うための基礎技術開発を目的としている点で新規性が高い。2つのタスクを設定しているが、それぞれ、商品の他商品と比較した場合の特長をレビューテキスト集合から自動抽出するタスク、二つの文書間の文レベルの対応を自動抽出するタスクであり、タスク設定自体が有意義でかつ新規である点が評価できる。

商品特長を自動抽出する研究では、対数尤度比と呼ばれる統計的な指標を用いることによって、商品特長を抽出できることを明らかにした。また、膨大なレビューテキストを有効に活用するために、特長を大きなカテゴリーに分類し、カテゴリー毎に情報を集約した統計量を用いる手法を考案した点が特に工夫した手法であると認められる。評価実験では、実際の電子商取引 web サイトに掲載されている数百万という大規模なレビューデータを用いて提案手法の有効性を実証しており、その結果の信頼性は高い。

二つの文書対間の文対応を抽出する研究では、次の点に新規性が認められる。(1)文種別を文対応推定の素性として利用することを提案。(2)Linear-chain/2D CRF による文対応推定モデルの提案。(3)文種別と文対応を同時に推定する統合モデルの提案。また、評価実験によって、単純な Logistic 回帰モデルなどによる文対応手法よりも高い精度で推定できることを示し、提案手法の有効性を実証している。

実用的な文書作成支援システムを構築するには、本論文で提案されている基盤技術で推定される情報をどのように提示するかなどのインターフェース的な追加研究が必要と思われる。しかし、新しい文書作成支援タスクを提案し、そのために必要な基盤技術を確立した点では意欲的かつ有用な研究であると評価できる。

【最終試験の結果】

平成 28 年 2 月 1 日、システム情報工学研究科において、学位論文審査委員の全員出席のもと、著者に論文について説明を求め、関連事項につき質疑応答を行った。その結果、学位論文審査委員全員によって、合格と判定された。

【結論】

上記の学位論文審査ならびに最終試験の結果に基づき、著者は博士（工学）の学位を受けるに十分な資格を有するものと認める。