

氏名	聶添		
学位の種類	博士(工学)		
学位記番号	博甲第9010号		
学位授与年月日	平成31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	ウェブ検索者の情報要求観点の集約に関する研究		
主査	筑波大学 教授	博士(工学)	宇津呂 武仁
副査	筑波大学 教授	博士(工学)	古賀 弘樹
副査	筑波大学 教授	博士(工学)	矢野 博明
副査	筑波大学 准教授	博士(情報科学)	星野 准一
		博士(デザイン学)	
副査	筑波大学 准教授	博士(工学)	乾 孝司

論文の要旨

本論文では、検索エンジン・サジェストを情報収集の情報源として、ウェブ検索者の情報要求観点を網羅的に収集・集約する手法の確立を目的とする。さらに、本論文では、日本と中国の関心事項の違いを分析することで、日中間文化間差異の発見を支援することをもう一つの目的として取り上げる。以上の背景をふまえて、本論文では、「ウェブ検索者の情報要求観点の網羅的収集・集約」、および、「日中間の文化間差異発見のための比較対照分析」の課題について取り組んでいる。

「ウェブ検索者の情報要求観点の網羅的収集・集約」の課題においては、検索エンジン・サジェストを情報要求観点ととらえて、その集約を行う先行研究のタスクにおいて、性能改善のための新たな知見を得ている。具体的には、従来より、収集されるサジェスト、および、ウェブページ集合の双方において、話題の内容が重複し冗長である点が問題であった。それを解消するために、文書集合にLDAトピックモデルを適用し、話題が類似するサジェストをまとめて俯瞰する手法が提案されていた。しかし、トピックモデルを用いた話題集約においては、集約の粒度が粗い点が問題であった。そこで、深層学習技術によりその問題を解消する方式を実現し、有効性を示している。

「日中間の文化間差異発見のための比較対照分析」の課題においては、先行研究において、トピックモデルにおけるトピックの単位での日中間文化間差異発見および比較対照分析における粒度が粗いという問題点を解消する方式に取り組んでいる。具体的には、トピック内の小分類単位を導入し、この単位での日中間比較対照分析方式を実現している。その結果、先行研究においては日中両側において観測されていた粗い事象に対し、より詳細な粒度の点で日中間差異が認められることを示している。

審査の要旨

【批評】

本論文で明らかにした事項は以下の通りである。

(1)「ウェブ検索者の情報要求観点の網羅的収集・集約」の課題においては、トピックモデルを用いた話題集約において、集約の粒度が粗い点が問題であった。そこで、この課題では、深層学習技術によりその問題を解消する方式を実現している。具体的には、近年、深層学習技術により、単語の意味表現を分散表現によって表す方式が提案され、その有効性が報告されている。分散表現方式においては、通常、大規模なコーパスから語の意味のベクトル表現を学習し、これを用いて各語の周囲の語の予測を行う。そこで、本論文では、分散表現によって表現される話題の粒度が相対的に細かいところに着目し、分散表現とトピックモデルを併用することによって、検索エンジン・サジェストの類似度を詳細に測定しトピックごとに細分化することによって、小分類の抽出を行っている。評価実験においては、人手によって作成された参照用小分類データに対して、3つの素性を設定し、二種類の分散表現訓練用コーパスを用いた場合について、小分類集約結果の適合率・再現率の比較を行っている。結果として、クエリ・フォーカス「就活」および「結婚」の両方において、分散表現訓練用コーパスとして「Wikipediaの全ページテキストのみ」を用いた場合より、「Wikipediaの全ページテキスト + 検索エンジン・サジェストによって収集されたウェブページテキスト」を用いた場合の方が、高い性能が得られることを示している。

(2)「日中間の文化間差異発見のための比較対照分析」の課題においては、トピックモデルにおけるトピックの単位での日中間文化間差異発見および比較対照分析において、比較対照分析の粒度が粗いという点が問題であった。そこで、この課題では、トピック内の小分類単位を導入し、この単位での日中間比較対照分析方式を実現している。具体的には、「日中共通で観測したトピック」に対して、その中に含まれる小分類を比較対照分析することで、「日本語側のみで観測した小分類」、「中国語側のみで観測した小分類」、「日中共通で観測した小分類」の三種類に分類できることを示している。結果として、「就活」の話題を対象とした日中間比較対照分析において、中国語側でのみ観測された小分類として転職活動に関する内容があるのに対して、日本語側でのみ観測した小分類として就活生が関心を持つ内容に関する小分類があった。これらの分析結果から、採用活動において重視する点の日中間の違いや両国の雇用形態の違いを発見することに成功している。

以上の議論に基づき、本論文は博士論文に値するものと認められる。

【最終試験の結果】

平成 31 年 2 月 13 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。