

氏名(本籍)	原田 智彦		
学位の種類	博士(工学)		
学位記番号	博 甲 第 7275 号		
学位授与年月日	平成 27 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	口語的な表現を含むテキスト情報の理解支援に関する研究		
主 査	筑波大学 教授	博士(工学)	津田 和彦
副 査	筑波大学 教授	博士(工学)	イリチュ 美佳
副 査	筑波大学 教授	博士(工学)	吉田 健一
副 査	筑波大学 准教授	博士(システムズマネジメント)	倉橋 節也
副 査	近畿大学 准教授	博士(工学)	溝渕 昭二

論 文 の 要 旨

近年のインターネットの普及に伴い、電子メールやFacebook、Twitterやブログ等のソーシャルメディア上には、日々、ユーザーが発信する膨大な量のテキスト情報がやり取りされている。通常、これらのテキスト情報の多くは、プレーン・テキスト形式であり、装飾などもない。さらに、最近では携帯電話やスマートフォンが普及し、キーによる文字入力や画面の狭小性のために長文を避ける傾向がある。また、匿名性の高さや比較的若い世代の利用者が多いため、記載されている文は、口語的な文体に近くだけた表現を多く含み、語句の省略も多い。

本研究では、上記の特徴を持つ電子メールやソーシャルメディア上に記載されたテキスト情報の正確な理解を支援することを目的とし、自然言語処理技術を用いた手法を提案する。具体的には、問い合わせメールに多く発生する省略の問題に着目し、省略された言葉を推定する方法を提案する。また、ソーシャルメディアについては、分析目的で収集したツイート中に分析に必要なない同じ名前の別の企業名や商品名が混入する問題(=同名他社の問題)に着目し、ツイート上の同名他社を判定する方法を提案する。

本論文は6章で構成される。1章はイントロダクションで、この研究の目的を概説している。2章では、口語的な表現を含むテキスト情報に起因する問題について説明すると共に、本研究で使用する自然言語処理分野における先進的な手法であるトピックモデルおよびLDA(Latent Dirichlet Allocation; 潜在的ディリクレ配分法)について解説している。次に3章では、前章で述べた課題を明確にするため予備実験を行っている。具体的には、口語調テキストと非口語調テキストでの解析結果の差異とその原因を明らかにすると共に、口語調テキストで頻発する主語などの名詞句が省略された場合の解析精度について実験結果にて確認している。4章では、本研究で提案する問い合わせメール中の省略を推定するアルゴリズムについて説明し、評価実験により提案したアルゴリズムの有効性

を確認すると共に考察を行っている。5章では、本研究で提案するツイート上の同名他社を判定する手法について説明し、評価実験により提案した手法の有効性を示すと共に考察を行っている。最終の6章では結論を示している。

審 査 の 要 旨

【批評】

インターネットの普及に伴い、電子メールやソーシャルメディア上には、ユーザーが発信した膨大なテキスト情報が存在する。これらの情報は、マーケティングや製品評判の把握など、企業にとっては極めて有益な情報であるが、多くは口語調で記述されており、その解析は現状の自然言語処理技術における意図理解では困難である。

本研究では、この口語調で記載されたテキストの解析精度の向上を目的に、省略語の推論アルゴリズムと、同音異義語の判定手法を提案している。省略語の推論アルゴリズムでは、N-gram モデルと LDA モデルの組み合わせによるスコアリングアルゴリズムを提案し、Yahoo!知恵袋のデータ約 57 万件を用いて評価実験を行い、その有効性を示している。また、同音異義語の判定では、ツイートの表層的な情報を利用した分類器と、ツイートを投稿したユーザーの興味モデルを使った分類器を基に、分類確率を線形補間したスコアリングモデルを提案し、1 万件のツイートにて実験を行い、その有効性を示している。これらの提案では、従来手法と比較して正解率を 10%以上向上させており、この結果は、自然言語処理の意図理解においては特筆に値する。

以上、社会人としての実務ノウハウと自然言語処理技術における意図理解の精度を向上した本学位論文、及び、研究の内容は、博士(工学)を授与するに十分なものと判断する。

【最終試験の結果】

平成27年1月27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。この結果とリスク工学専攻における達成度評価による結果に基づき、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。