

2023年2月10日

豊橋技術科学大学長 殿

情報・知能工学専攻

学位審査委員会



委員長

岡田 美智男



論文審査及び最終試験の結果報告

このことについて、博士学位論文審査を実施し、下記の結果を得ましたので報告いたします。

学位申請者	Setio Basuki		学籍番号	第199304号
申請学位	博士 (工学)	専攻名	大学院工学研究科博士後期課程 情報・知能工学 専攻	
博士学位論文名	The Crucial Role of Citation Functions in The Technology-assisted Peer Review (計算機によって支援されたピアレビューにおける引用目的の役割に関する研究)			
論文審査の期間	2023年1月12日 ~ 2023年2月10日			
公開審査会の日	2023年2月1日	最終試験の実施日	2023年2月1日	
論文審査の結果※	合格	最終試験の結果※	合格	
審査委員会(学位規程第6条)				
学位申請者にかかる博士学位論文について、論文審査、公開審査会及び最終試験を行い、別紙論文内容の要旨及び審査結果の要旨のとおり確認したので、学位審査委員会に報告します。				
委員長	北岡 教英			
委員	梅村 恭司		印	
	土屋 雅稔		印	
		印	印	

※論文審査の結果及び最終試験の結果は「合格」又は「不合格」の評語で記入すること。

論文内容の要旨

本論文は、科学論文の査読作業を支援する技術を開発することを目標として、科学論文における先行研究の引用目的の分類とデータベース化を行い、計算機科学分野の国際会議論文を対象として、先行研究の引用目的と論文の品質および査読スコアの関係を検証するものである。

第1章は、科学論文の査読作業を支援する技術の必要性和現状、および先行研究の適切な引用が査読結果に対して及ぼす影響について概観した。第2章は、科学論文における先行研究の引用目的を分類する体系を提案した。この体系は、5種類の大分類と21種類の小分類からなる2段階の体系である。作業者によるアノテーション実験では、大分類レベルで0.85、小分類レベルで0.71という十分に高い一致度が得られた。第3章では、提案した体系に基づき、Active Learningを適用して大規模データセットを半自動的に構築した。構築したデータセットは、計算機科学分野の国際会議論文を対象として1,840,815事例からなり、引用目的に関するデータセットとして最大規模のデータセットである。第4章では、第3章のデータセットから学習した引用目的分類モデルを用いて抽出した特徴量に基づく機械学習モデルにより、計算機科学分野の国際会議論文を対象として、論文の品質、採択結果および査読スコアを推定する方法を提案した。第5章では、COVID-19に関して発表された論文を対象として引用目的データセットを作成し、計算機科学分野との違いを明らかにした。第6章では、本論文の限界について言及した上で、本論文を総括して今後の展望を述べている。

審査結果の要旨

近年、国際会議や雑誌に投稿される論文の増加にともなって、査読者に対する負担が急増しており、査読システムの継続性が問題となっている。その解決策として、3つの方向が検討されている。第1は、適切な分散化など査読システムの調整によって、査読者の負担を軽減する方向である。第2の方向は、査読者の負担を軽減する技術的な手段の提供である。第3の方向は、査読を必要としない科学研究体制の構築である。現状では、第1の方向は既に実践例があるが、第2の方向の実践例は存在しない。本論文は、第2の方向で査読者の負担を軽減することを目標として、先行研究の引用目的の分類とデータベース化を行い、先行研究の引用目的と論文の品質および査読スコアの関係を検証するものである。

第2章では、科学論文における先行研究の引用目的を分類する体系を提案している。この体系は、トップダウンなアプローチとボトムアップなアプローチを併用して、既存の体系よりも広い分野の論文を考慮している点が新しい。第3章では、Active Learningを利用して大規模な引用目的データセットを半自動的に構築する方法を提案している。提案手法は、非常に大規模なデータセットを比較的低コストに実現している点に意義がある。第4章は、第3章のデータセットから学習した引用目的分類モデルを用いて抽出した特徴量に基づいて、論文の品質、採択結果および査読スコアを推定する方法を提案している。この方法は、査読者の負担を軽減するという目標を考慮して、査読スコアや査読コメントを参照していない点が既存研究とは異なる。加えて、第5章では、COVID-19に関する科学論文を対象とするデータセットの構築を通じて、本論文の提案手法が、計算機科学分野以外にも適用できることを示している。また、2つのデータセットの比較を通じて、計算機科学分野の論文では、先行研究と性能面で比較している引用の存在が、論文の品質や採択結果の推定にあたって重要であることを示している。

本論文は、2つの明らかな限界を持つ。第1に、論文の品質や採択結果、査読スコアの推定結果に関する根拠の説明が不十分である。第2に、他の研究分野に対しても同様に適用できるか不明である。このような限界は存在するものの、引用目的の分類体系の提案、大規模な引用目的データセットの構築方法、および引用目的に基づいて論文の品質を推定する方法を見出していることから、本論文は博士（工学）の学位論文に相当するものと判定した。

(各要旨は1ページ以上可)