

引用コンテキストに基づく  
高影響論文の推定に関する研究  
A Study on Estimation of High Impact Paper  
based on Citation Context

学籍番号：201721659

氏名：神澤 義人

Yoshito Kamisawa

この論文は科学文献の引用構造に焦点を当てている。学術文献における引用の目的は、その論文が前提とする研究や、モチベーションとなった技術、提案手法の比較対象などを明らかにすることであり、一般的に複数の論文を列挙している。一方、学会や雑誌で発表された学術論文そのものに着目すると、論文の本文はテキスト基軸として章立てされており、研究成果を理論的に説明するための構造化がなされている。

本論文では、本文中に出現する引用マークを手がかりとして、構造化された文献の引用箇所を特定し、引用箇所に基づいて引用の役割をラベル付けする。被引用文献に付与するラベルは、A:abstract, I:Introduction, R:relatedWorks, F:FutureWorks, C:Conclusion, M:Methodsの6種類である。これにより、多くの学術文献から引用される文献は、多くのラベルを付与されることになる。このラベル集合と引用されるまでの期間を特徴量として、高影響論文を推定する手法を提案する。ここで、高影響論文とは、その後の研究に大きな影響を及ぼした論文を言う。提案する特徴量の有効性を検証するために、高影響論文とそうでないものを提案する特徴量で分離できるかを実験した。実験の対象とした文献は、データベース分野で最重要会議として位置づけられているVLDB(Very Large Data Base Endowment Inc.)による国際会議と予稿集で発表された35年分、4,412件である。VLDB Conferenceでは、論文公開から10年後に判断する10-Year Awards賞を設けていることから、当該賞を受賞した論文を高影響論文とみなし、受賞しなかった論文と分類できるかを実験した。

高影響論文を判別する引用の特徴が明らかになることで、これまで多大な労力を要していたAward論文の選別作業が軽減できる、比較的短期間の引用傾向を観測することで高影響論文の候補を知ることができるなどの効果が期待される。

研究指導教員：佐藤 哲司

副研究指導教員：若林 啓