

학술문헌을 인용하는 트윗의 기능 분석 연구

김병규*, 강지훈^o

*한국과학기술정보연구원 데이터기반문제해결연구단

^o충남대학교 컴퓨터공학과

e-mail: bk.kim@kisti.re.kr*, jhkang@cnu.re.kr^o

Function Classification of tweets Citing Scholarly Articles

Byungkyu Kim*, Ji-Hoon Kang^o

*Dept. of data-centric problem solving research, KISTI

^oDept. of Computer Engineering, Chungnam National University

● 요약 ●

개별논문 평가를 위해 제안된 altmetric가 주목받고 있다. altmetrics에서는 개별 논문의 트윗의 건수를 평가요소 중 하나로 활용한다. 그러나 여러가지 목적으로 작성된 트윗을 단일하게 처리하는 것은 문제가 있다. 본 논문은 과학 논문에 달린 트윗들을 분석하여 기능의 범주를 정의하고 분류체계를 제시하였으며, 기존의 논문의 인용기능 분류 실험을 실시하여 그 결과와 비교 분석을 수행하였다. 향후 도출한 트윗 기능 분류에 대한 개선과 추가적인 연구를 수행할 계획이다.

키워드: 알트메트릭스(Altmetrics), 인용기능(Citation Function), 소셜네트워크서비스(SNS), 트윗(Tweets)

I. Introduction

과학기술정보(논문)의 생산규모가 빠르게 증가하고 있으며 유통의 주기도 갈수록 빨라지고 있다. 기존의 동료평가 및 인용분석은 그 특성상 연구결과에 대한 평가가 매우 느린 특성으로 현재와 같은 빠른 과학기술정보의 순환주기에서 빠르고 정확한 추세 분석에 한계가 있다. 이에 보다 빠른 과학기술정보 탐색과 분석을 위하여 소셜 네트워크 서비스(SNS)를 활용한 과학기술정보 탐색 및 분석과 관련된 해외의 연구가 집중하고 있다. 논문단위의 새로운 연구성과 영향력 측정지수인 알트메트릭스는 기존의 논문들 사이의 인용 횟수 뿐만 아니라 소셜미디어에서의 공유 및 언급 횟수 등까지를 측정 범위로 삼으며 현재 Altmetrics, PlumX 등의 서비스들이 개발되어 제공되고 있다. 알트메트릭스에서는 개별 논문의 트윗의 건수를 평가요소 중 하나로 활용한다. 그러나 여러가지 목적으로 작성된 트윗을 단일하게 처리하는 것은 문제가 있다. 본 논문에서는 과학 논문에 달린 트윗들을 분석하여 기능의 범주를 정의하고 분류스킴(체계)을 제시한다. 또한 기존의 논문의 인용기능 분류 실험을 실시하여 그 결과와 비교 제시하였다.

대상으로 선정하였다. Altmetric사는 PLOS 학술논문 데이터셋을 대상으로 Altmetric을 적용한 데모사이트를 개발하여 제공하고 있는 점도 PLOS를 선정한 배경이 되었다. PLoS Impact Explorer는 altmetric.com에서 수집된 Public Library of Science (PLoS) 출판논문에 대한 Altmetric 정보를 서비스 한다. 선정된 수집처에서 - Altmetrics점수가 가장 높은 논문 85개를 데모사이트를 통해 수집하였다. 사이트내 검색 조건은 1. all time, 2. PLoS에서 발행하는 모든 주제분야 논문(총 5,166건)을 대상으로 하였다.

• 데이터 필터링 방법

- 1) 저널 및 논문제한 : PLOS 논문 85개 중 분석을 위한 이해가 상대적으로 용이한 PLOS ONE 저널 논문 25개로 제한
- 2) 소셜미디어 소스제한 : 소셜미디어 중 가장 학술논문에 대해 인용이 활발한 “트위터”로 제한
- 3) 최종분석 대상 선정 : PLOS One 25개 논문의 “트윗” 레코드 6,683건 중에서 비영어/단순링크/단순제목 기술의 경우 제외한 2,521건으로 분석대상 확정

II. Materials and Methods

소셜미디어를 통한 학술문헌 인용사례를 분석하기 위하여 Altmetrics를 직접 개발하고 학술정보서비스에 적용하는 PLOS를

II. Result

2장에서 수집된 논문에대한 트윗들에 대한 기능들을 분석하여

세부 기능 분류표를 정의하였다. 도출된 기능으로는 “내용기술” “동의” “비판” “제안” “추천” 등이었으며 학술논문과 같이 정제된 문헌에서의 인용이 아니므로 트윗의 내용에 대한 기능은 기존 인용문 기능 분류들보다 매우 다양하게 분석되었다. 학술문헌에 대한 트윗 분석 결과는 아래 그림1과 같다. 내용기술이 가장 많은 비율을 차지하며 동의와 흥미, 추천 등이 그 다음 순으로 분석되었다.

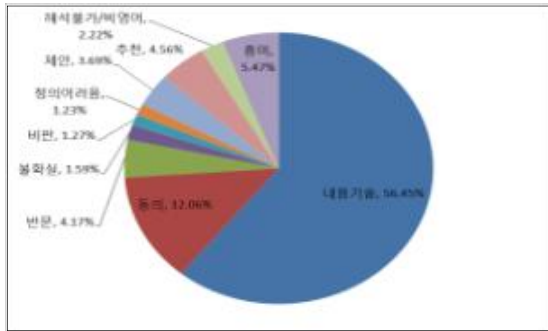


Fig. 1. Result of Function Classification of Tweets Citing Collected Articles

Table 1. Citation Function Classification Sample

인용기능	설명
COCOGM	방법론과 연구목표에서의 대조
COCORo	결론의 중립적인 비교
COCOXY	다른 두 인용 대조
Neut	중립
PMOT	문제동기화 및 해결책으로 선택
Psim	접근방법 또는 목표의 유사성
Psup	지지 또는 호환/서로지지
Puse	일부 측면의 사용

학술논문의 인용문에 대한 인용기능에 대한 일반적인 기능 분류의 예는 아래 표1과 같다. 인용문헌에 대한 사용, 지지, 호환, 대조 등 트윗 기능보다 매우 구체적인 기능으로 분류될 수 있음을 보여준다.

2장에서 수집된 논문들에 대한 학술논문들에서의 인용문을 표1의 인용기능들을 사용하여 분류하면 아래 그림2와 같다. 일부측면의 사용이 가장 많은 비율을 보였으며, 중립과 지지가 그 다음 순위로 분석되었다.

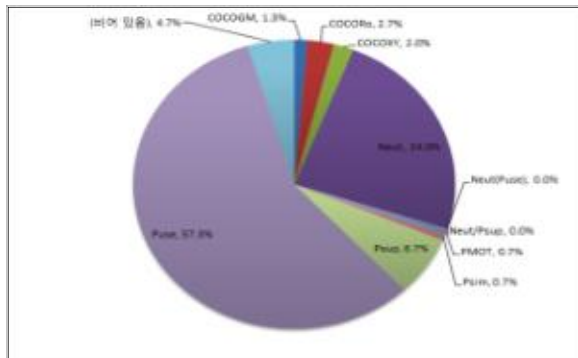


Fig. 2. Result of Function Classification of Citing Articles Citation Sentence

일부 측면의 2장에서 수집된 논문들에 대한 학술논문들에서의 인용문을 표1의 인용기능들을 사용하여 분류하면 아래 그림2와 같다.

II. Conclusion

본 논문에서는 과학 논문에 달린 트윗들을 분석하여 기능의 범주를 정의하고 분류스킴(체계)을 제시하였다. 또한 기존의 논문의 인용기능 분류 실험을 실시하여 그 결과와 비교하였다. 구체적으로 학술문헌의 트윗에 대한 분류를 진행한 결과 전통적인 학술문헌의 인용분류표와는 상이한 분류결과가 도출되었다. 이는 정형화된 포맷으로 작성되는 정형화된 학술출판물과 자유롭게 의견을 개진하는 트위터의 근본에서 비롯되는 것으로 보여진다. 향후 본 연구에서 제시한 트윗의 기능분류를 개선하고 추가적인 분석 연구를 진행할 계획이다.

REFERENCES

- [1] Altmetrics Site, <http://altmetrics.org>
- [2] Plos (Public Library of Science), <http://www.plos.org>
- [3] Altmetric Service, <http://www.etric.com>
- [4] PlumX Service, <https://plumanalytics.com/learn/about-metrics/>