
뉴스 데이터 분석을 통한 미래 정보통신의 주요 기술 탐색

이상규

한국건설기술연구원

Searching for New Challenge of Information and Communication Technology in News Articles with Data Analysis

Sanggyu Lee

Korea Institute of Civil Engineering and Building Technology

E-mail : sanggyulee@kict.re.kr

요 약

최근, 데이터 분석을 통해 새로운 이슈를 선제적으로 확인하고 이를 통해 기술 발전을 이끌어 내고 있다. 이러한 새로운 이슈에 대해서 미디어 보도가 결정적인 역할을 하며, 특히 과학 기술에 대한 사회적 인식 형성에 큰 영향을 끼치고 있다. 이에 따라 200여개의 뉴스 기사를 중심으로 키워드 분석과 감정 분석, 2가지 데이터 분석을 바탕으로 미래 정보통신의 주요 기술(Machine Learning & Blockchains)을 확인한다. 본 분석 결과를 바탕으로 향후 정보통신 주요 기술을 지속적으로 발전하는데 길잡이 역할을 할 것으로 예상된다.

ABSTRACT

Recently, people are using the data analysis in order to follow the new trend in information and communication technology. Media plays an important role to expand the new issue in our society, especially affected to establish social awareness about science and technology. So, We find some major technologies (Machine Learning & Blockchains) of future communication and information based on the 200 news articles through two data analysis methods such as keyword analysis and sentiment analysis. We look forward this paper to constantly develop the technology of information and communication as the guiding frame of the new scientific world.

키워드

Machine Learning, Blockchains, News Articles, Keyword Analysis, Sentiment Analysis

1. 서 론

“이전과는 전혀 다른 기술혁명이 점차 다가오고 있다. 이러한 기술혁명은 규모와 범위 등 전혀 다른 차원의 방식으로 경험하게 될 것이다“

이는 2016년 다보스포럼 중 세계경제포럼 회장인 클라우드 슈밥이 미래 기술에 대해 이야기하면서 제4차 산업혁명의 개념에 대해 처음 언급하였다.

현재 우리가 4차 산업혁명 중 가장 이슈화가 되고 있는 기술이 바로 정보통신간의 융합기술이다. 예를 들면, IoT, Cloud Computing 등 다양한 정보통신기술이 있다. 특히 이러한 기술을 지속적

으로 분석하고 제공하는 곳 중이 바로 세계적 IT 자문기관인 가트너(Gartner)이다. 매년 가트너가 선정한 10대 전략 기술 트렌드는 향후 산업 트렌드에 지대한 영향을 미치고 있기에 정부, 민간 등 다양한 기관에서 매년 가트너의 10대 기술을 주목하고 있다[1].

가트너에 따르면, 2017년 10대 전략기술은 아래 표 1와 같이 정리하였다.

표 1. 가트너가 선정한 2017년 10대 전략기술 트렌드

	기술명	구분(영역)
1	Applied AI & Advanced Machine Learning	Intelligent
2	Intelligent Apps	
3	Intelligent Things	
4	Virtual Reality and Augmented Reality	Digital
5	Digital Twins	
6	Blockchains and Distributed Ledgers	
7	Conversational Systems	Mesh
8	Mesh App and Service Architecture	
9	Digital Technology Platforms	
10	Adaptive Security Architecture	

제시된 표 1에 나타난 10대 전략기술 중 정보통신기술에서 가장 많이 이슈화 되고 있는 전략기술 2개(Machine Learning & Blockchains)를 뽑아 해당 키워드를 바탕으로 뉴스 기사 정보를 수집하였다.

II. Data Collection

한 연구결과에 따르면, 대중미디어는 과학기술에 대한 대중의 이해와 태도에 영향을 주고 있다고 한다[2]. 따라서 본 연구 데이터를 위해 2016년 7월 이후에 나온 웹 뉴스 상 Machine Learning과 Blockchains 2개의 키워드를 중심으로 각 200여개의 뉴스 기사를 검색하였고 이를 바탕으로 해당 뉴스 데이터를 기초로 Database를 확보하였다.

과학기술에 대한 대중의 감성을 조성하기 위해서는 미디어가 매우 중요한 역할을 담당하고 있다는 기존 연구 결과에 따라, 뉴스데이터 기반 감성분석을 통해 각 정보통신 기술에 관한 흐름을 파악한다[3].

단, 해당 데이터에 대한 분석을 진행하기 앞서 뉴스는 주로 중립성을 띄는 기사로 구성되어 있기 때문에 해당 키워드에 대한 주요 시사점 및 요약문을 위주로 데이터를 분석하였다.



그림 1. 데이터 수집 및 구축 과정

본 연구의 데이터 수집 및 구축과정 모형을 위 그림 1을 통해 제시하였다.

III. Data Analysis

(1) 감성분석

감성분석 연구는 자연어 처리를 바탕으로 다양한 분야에서 활발히 연구되고 있는 분석방법이다. 특히, 비즈니스 측면에서 대중이 상품 또는 서비스에 대해서 어떻게 생각하는지를 이해함으로써 경영진의 의사결정에 큰 영향을 미치기도 한다.

본 연구에서는 R (Version 3.3) 프로그래밍을 바탕으로 부정단어와 긍정단어를 점수화하여 긍정과 부정 형태의 빈도 분석을 확인하였다.

1) Machine Learning

우선, 2017년도 가트너 10대 기술 중 첫 번째로 채택이 된 Machine Learning을 차세대 정보통신 기술 키워드의 하나로 선택하였다.

R을 바탕으로 200개의 뉴스데이터를 감성 분석한 결과, 109개의 긍정반응과 33개의 부정반응, 그리고 중립성 반응이 56개로 확인되었다.

지난 2010년부터 꾸준히 부각되고 있는 Machine Learning 기술은 이미 많은 산업 분야에 활용하고 있는 기술이기 때문에 대중미디어로부터 매우 호의적인 반응을 보이고 있다고 판단할 수 있다.

2) Blockchains

2017년도 가트너 10대 기술 중 여섯 번째로 채택이 된 Blockchains을 차세대 정보통신 기술 키워드의 하나로 선택하였다.

R을 바탕으로 200개의 뉴스데이터를 감성 분석한 결과, 61개의 긍정반응과 19개의 부정반응, 그리고 중립성 반응이 120개로 가장 많은 데이터를 확보하였다.

참고문헌

- [1] Kasey Panetta "Gartner's Top 10 Strategic Technology Trends for 2017" Gartner, 2016년
- [2] 조항민, "국내미디어 보도에 나타난 생명공학," BioSafety 바이오세이프티, 제8권, 제1호, pp.32-43, 2007년.
- [3] 이용섭, 문영규, 서승현, "감성과 과학기술 위험수용성의 동기화 연구," GRI 연구논총, 제14권, 제3호, pp.217-240, 2012년.