

텍스트 마이닝을 활용한 4M 기법 기반의 건설사고 요인 분석

강동욱*, 차기춘**, 조민건***, 박승희****

*성균관대학교 글로벌스마트시티융합전공 석사과정

e-mail : kdu7964@g.skku.edu

**성균관대학교 글로벌스마트시티융합전공 연구교수, 공학박사

e-mail : ckckicun@skku.edu

***성균관대학교 글로벌스마트시티융합전공 석사과정

e-mail : raonik6713@naver.com

****교신저자, 성균관대학교 건설환경공학부 교수, 공학박사

e-mail : shparkpc@skku.edu

Analysis of Construction Accidents Factors Based on 4M Method Using Text Mining

Donguk Kang^{*}, Gichun Cha^{**}, Mingeon Cho^{***}, Seunghee Park^{****}

^{*, **, ***}Dept. of Global Smart City, Sungkyunkwan University

^{****}School of Civil, Architectural & Environmental Engineering, Sungkyunkwan University

요약

국내 건설현장에서는 사망자가 전체 산업재해 중 50% 이상 차지하고 있으며, 여전히 인명피해는 줄어들지 않고 있다. 건설현장에서 작업자의 안전이 최우선으로 보장되어야 하지만, 실제 현장에서는 작업자도 확인할 수 없는 위험요인들이 다수 존재한다. 이에 따라 건설현장의 건설사고 요인을 분석하여 작업자의 안전을 보장할 수 있는 정보를 마련하고자 한다. 본 논문에서는 텍스트 마이닝을 기반으로 건설현장 사고 데이터를 활용하여 건설사고 요인을 분석하였다. 건설사고 데이터는 건설공사 안전관리 종합정보망(CIS)에서 수집하였으며, 텍스트 마이닝을 통해 분류된 데이터를 바탕으로 건설사고사례를 4가지 측면으로 구분하고, 건설사고의 원인과 결과를 도출하였다. 텍스트 마이닝을 활용하여 인과 네트워크를 구축하고 분석한 문헌 검토 및 자연어 프로세스(NLP)기법을 적용하였으며, 4M 기반의 건설사고 요인 인자를 추출하였다. 추출된 건설사고 요인은 건설사고 예방을 위한 자료로 활용될 수 있으며, 작업자의 인명피해를 줄일 수 있을 것으로 기대한다.

1. 서론

대한민국 고용노동부에 따르면 건설현장에서는 2015년부터 2019년까지 5년간 약 2만5000명의 건설노동자가 산업재해로 피해를 보았다. 또한, 이전 시대에 걸쳐 매년 약 500명의 사망자가 발생하고 있다. 2020년 산업재해 사고사망자 882명 중(2020년 4월 14일 기준) 건설업 사고사망자는 458명(51.9%)으로 전년 대비 30명 증가했고 사고사망 만인율 또한 전년 대비 0.28% 상승한 2.00%로 나타났다(고용노동부, 2020). 기술 발전에도 불구하고, 건설에서의 안전관리는 여전히 중요한 과제로 남아있다.

건설현장에서 발생하는 사고는 치명적인 부상과 사망을 초래하여 기업은 물론 개인에게 경제적 손실과 사회적으로 부정적인 영향을 초래한다 [1].

국도교통부에서는 건설공사 안전관리 종합정보망(CSI)를 통해 건설공사에 발생한 산업재해를 건설사고 사례 보고서로 공지하고 있으며 텍스트 데이터로 설명이 제공되어 비정형 데이터 분석에 활용되었다 [2].

건설사고사례 보고서 분석을 위해 텍스트 마이닝을 활용한 연구를 통해 건설현장 사고를 분석하였다. 이러한 분석을 인적, 물질적, 환경적, 관리적 4M 측면으로 분류하여 높은 위험요인과 낮은 위험요인을 Wordcloud에 적용 후 건설사고의 중요도에 따른 주요 요인 인자를 시각적으로 표현하였다 [3].

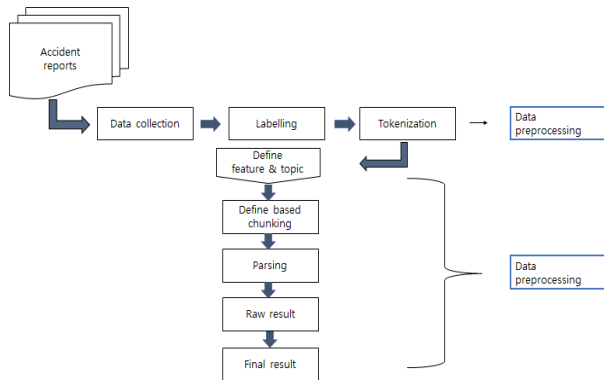
2. 본론

2.1 데이터 수집 및 분류

2.1.1 데이터 수집

건설사고사례 보고서 분석을 위해 건설공사 안전관리 종합

정보망(CSI)에서 웹 크롤링 과정으로 최근 약 2년간의 4518건의 사고 데이터와 15개의 유형으로 수집하였다. 본 연구에서는 그림1과 같이 건설사고사례에 대한 비정형 텍스트 데이터를 활용하여 규칙을 기반으로 데이터를 저장하였다.



[그림 1] 데이터 수집을 통한 텍스트마이닝 분석 흐름도

2.1.2 데이터 분류

데이터 분류는 사고 데이터의 내용에 따라 유형으로 저장하는 과정에 사용되었다. 데이터 분류 작업은 사고에 대한 원인분석과 사고 경위, 사고원인, 피해 내용 등으로 구분되어 비정형 데이터를 활용하는 데 쓰였다. 이러한 데이터 분류로 인해 건설사고사례 보고서 분석에 대한 텍스트 마이닝에 필요한 데이터 분류 과정을 수행하였다 [4].

2.2 연구 방법

2.2.1 텍스트마이닝 기법을 적용한 사고분석

텍스트마이닝은 정의되지 않은 텍스트 데이터로부터 정보를 도출하는 프로세스로 정의된다. 여기에는 텍스트를 데이터 마이닝 알고리즘에 사용할 수 있는 숫자 데이터로 변환하는 작업이 포함된다 [5]. 본 연구에서 텍스트마이닝은 4M 기반으로 분류된 텍스트의 사고 요인 인자를 분석하는 데 사용되었다. 15개의 유형으로 나누어진 텍스트 데이터를 인적, 물질적, 환경적, 관리적 4가지 측면으로 키워드 간 연관 관계를 분석하여 향후 건설사고의 대책 수립에 반영될 수 있도록 연구를 수행되었다.

2.2.2 4M 기반 사고 요인에 미치는 영향

4M 기법은 예측 가능한 산업재해 발생 위험요인을 분석하여 사고 발생 가능성을 최소화하기 위한 것으로 위험요인을 Machine(기계적), Media(물질 및 환경적), Man(인적), Management(관리적)의 4가지 요인을 통해 구분하여 위험 제거 대책을 제시하는 방법이다 [6].

본 연구에서는 텍스트 마이닝 기법을 기반으로 분류된 데이터를 활용하여 4M 기법의 위험요인에 미치는 영향을 분석

하였다. 이후 사고 요인에 미치는 영향을 시각화하여 추출하였다.

3. 결론

본 연구에서는 텍스트마이닝을 활용하여 건설사고사례를 분석하여 원인과 결과를 도출하였다. 분석 결과 건설사고 원인은 4M 기법으로 구분할 수 있으며, 작업자의 안전을 관리·예측하여 건설사고 예방에 근거를 마련할 수 있을 것이라 기대한다.

감사의 글

이 연구는 국토교통부/국토교통과학기술진흥원이 시행하고 한국도로공사가 총괄하는 “스마트건설기술개발 국가R&D 사업(과제번호 21SMIP-A158708-02)”의 지원으로 수행되었으며, 국토교통부의 스마트시티 혁신인재육성사업으로 지원되었습니다.

참고문헌

- [1] Na XU, Ling MA, Qing Liu, Li WANG, Yongliang Deng. (2021), “An improved text mining approach to extract safety risk factors from construction accident reports”, Safety Science, Vol. 138, Article 105216.
- [2] 엄진기. (2019), “철도사고 데이터를 활용한 텍스트마이닝 분석”. 한국도시철도학회논문집, 제 7권 3호, pp. 397-405.
- [3] 김규진, 최병선, 전재역. (2016), “플랜트 시설물 시공단계의 4M기법을 활용한 전문가 인식조사에 의한 안전관리 요인분석”, 한국건설관리학회논문집, 제 17권 1호, pp. 18-27.
- [4] Botao Zhong, Xing Pan, Peter E.D.Love, Jun Sun, Chanjuan Tao. (2020), “Hazard analysis: A deep learning and text mining framework for accident prevention”, Advanced Engineering Informatics, Vol. 46, Article. 101152.
- [5] Fan Zhang, Hasan Fleyeh, Xinru Wang, Minghui Lu. (2019), “Construction site accident analysis using text mining and natural language processing techniques”, Automation in Construction, Vol. 99, pp. 238-248.
- [6] Ruipeong Tong, Hui Zhao, Na Zhang, Hongwei Li, Xiaolong Wang, Hongqing Yang. (2021), “Modified accident causation model for highway construction accidents (ACM-HC)”, Engineering, Construction and Architectural Management, Vol. 28, Issue 9.