

토픽모델링을 활용한 부산항 항만안전성 이슈 동향에 관한 연구

이정민* · 하도연** · † 김율성

*한국해양대학교 대학원 박사과정생, **한국해양대학교 대학원 석사과정생, † 한국해양대학교 물류시스템공학과 교수

요 약 : 최근 들어, 현대사회는 예측이 불가능한 다양한 위험성들이 존재하여 글로벌 의존도가 높은 항만물류산업의 위험부담이 증가하고 있다. 이에 본 연구에서는 항만산업의 안전성에 영향을 미치는 요인을 알아보기 위해 과거부터 현재까지 국내 항만 안전성에 영향을 미친 이슈들을 시계열적으로 살펴보고자 하였다. 이를 위하여 국내를 대표하는 부산항의 항만 안전성과 관련된 뉴스 기사 텍스트 데이터를 활용하여 LDA 토픽모델링 분석을 진행하여 부산항 항만안전 주요 이슈들의 동향을 살펴보고자 하였다.

핵심용어 : 항만안전, 부산항, 텍스트마이닝, LDA토픽모델링

1. 서 론

항만은 육·해상을 연결하는 중요한 연결공간이며 국가 수출입 물류 활동을 담당하는 공간이다. 항만물류산업은 국가 경제를 뒷받침하며 이러한 항만 기능이 마비되면 국가 경제와 국민 생활에 상당한 영향을 미칠 수 있다. 항만산업은 외부적인 영향으로부터 가장 먼저 민감하게 반응하므로 현대 사회에 산재하여 있는 글로벌 불확실성의 위험에서 안전할 수 없다. 최근에는 코로나19 확산, 러우전쟁 등 글로벌 공급망에 예측이 어려운 다양한 위험성이 존재하여 세계적인 의존도가 높은 항만물류산업의 위험부담이 증가하고 있다. 따라서, 국가 중요 시설인 항만의 안전성에 영향을 미치는 요인들을 알아볼 필요가(중략).....

2. 이론적 배경 및 선행연구 고찰

2.1 항만안전의 중요성

통합방위법 제2조 13호에 따르면, 항만은 국가 경제를 뒷받침하는 중요 국가기간산업이며, 외부의 적(영향)에 의해 그 기능이 마비될 경우 국가 안보 및 국민생활에 심대한 영향을 미치는 시설이다.

이러한 국가 중요시설인 항만의 환경은 광범위한 경제적, 사회적, 환경적, 제도적, 이슈와 충격들에 영향을 받으며.....(중략).....

2.2 선행연구 고찰

2.2.1 항만안전 관련 선행연구

Lam(2015)는 전 세계 공급망의 핵심 연결점 역할을 하는

항만의 운영 중단은 글로벌 무역망 흐름과 다양한 이해관계자에게 영향을 준다고 밝혔으며 항만산업과 해운산업이 글로벌 공급망의 불확실성을 초래하는 가장 주요한 원인이라고 밝혔다. 현재에는 기후변화, 국가적 안보, 기름유출, 사회정치적 불안 등으로 인한 우려가 커지고 있어 항만의 리스크를 더욱 높일 수 있으므로 항만의 운영에 대한 세심한 주의가 필요한 시점.....(중략).....

2.2.2 텍스트마이닝 관련 선행연구

Rha(2022)는 코로나19 팬데믹으로 인한 글로벌 공급망 불확실성으로 인해 초래된 컨테이너 운임 상승은 공컨테이너 부족과 재난사고, 항만폐쇄, 항만혼잡도의 악화, 등의 파급성 영향력으로 인해 상승세가 지속되고 있다고 밝혔다. 따라서 본 연구에서는 컨테이너선과 관련된 뉴스들을 대상으로 텍스트 네트워크 분석과 토픽모델링 분석을 활용하여 최근의 컨테이너 운임 상승세의 요인을 파악하고자 하였다. 본 연구에서는 해운 전문지인 로이즈리스트에 2020년 1월부터 2021년 7월까지 게재된 뉴스 기사들을 수집하여 분석하였으며 넷마이너 4.0을 활용하였다. 본 연구의 분석 결과는.....(중략).....

3. 연구설계 및 연구방법

3.1 연구설계

3.1.1 분석 개요

본 연구는 부산항 항만 안전성과 연관된 뉴스 기사 텍스트 데이터를 기반으로 LDA 토픽모델링 분석을 활용해 부산항 항만 안전성과 관련된 주요 이슈 동향을 5년 주기별로 알아보고자 하였다. (중략)

† 교신저자 : 정희원, logikys@kmou.ac.kr 051)410-4332

* 정희원, jmjm3646@g.kmou.ac.kr 051)410-4890

** 정희원, ehdu16091@g.kmou.ac.kr 051)410-4890

3.1.2 데이터 수집 및 전처리

본 연구에서는 빅카인즈에서 제공하는 부산항의 항만 안전 성과 연관되며 키워드에 “안전”, “위험”, “사고” 등의 단어를 포함하는 뉴스 기사를 추출하여 활용하였다. 수집된 뉴스 기사 데이터는 1993년 1월 1일부터 2022년 12월 31일까지 총 30년 동안 게재된 자료들이며 최종 수집된 뉴스 기사 데이터 건수는 총 3,929건이다. 본 연구에서는 30년을 5년 주기로 나누어 1기~6기로 살펴보았다. 수집된 데이터의 현황은 아래의 Table. 1과 같다. …… (중략) …….

Table. 2 Data Source

type	survey year	survey media	total
News	1993-2022	Big kinds	3,929

3.2 연구방법

3.2.1 LDA 토픽모델링

토픽모델링 분석이란 비정형의 텍스트 데이터에서 등장한 주제어들의 동시에 사용된 패턴을 바탕으로 해당 텍스트들을 대표할 수 있는 특정한 주제나 이슈들을 자동 추출하는 분석 기법이다. 이 분석기법은 텍스트 데이터 속에 잠재된 중요한 주제들의 도출 및 발견과 이들의 상호연결성, 그리고 시계열에 따른 토픽 변화를 발견할 수 있다.(Blei, 2012; Jung, 2022; Jang et al, 2023). 이때 등장하는 토픽은…… (중략) …….

4. 연구결과

4.1 부산항 항만안전 이슈 토픽모델링

4.1.1 1기(1993년-1997년)

1993년부터 1997년까지의 기간에 해당하는 1기동안 수집된 뉴스 기사는 총 69개이며 전처리 작업을 마친 단어의 개수는 총 3,814개이다. 이 데이터를 바탕으로 LDA 토픽모델링을 진행하였다. ……(중략)…….

4.1.2 2기(1998년-2002년)

1998년부터 2002년까지의 기간에 해당하는 2기동안 수집된 뉴스 기사는 총 71개이며 전처리 작업을 마친 단어의 개수는 총 3,128개이다. 이 데이터를 바탕으로 LDA 토픽모델링을 진행하였다. ……(중략)…….

4.1.3 3기(2003년-2007년)

2003년부터 2007년까지의 기간에 해당하는 3기동안 수집된 뉴스 기사는 총 167개이며 전처리 작업을 마친 단어의 개수는 총 3,429개이다. 이 데이터를 바탕으로 LDA 토픽모델링을 진

행하였다. ……(중략)…….

…… (중략) ……

5. 결론

본 연구는 지난 30년(1993년-2022년)간 부산항 항만안전과 관련된 국내 뉴스 기사 3,929건의 텍스트 데이터를 토픽모델링과 시계열 회귀분석을 이용하여 부산항 항만안전 주요 이슈를 유형화하고 이슈 동향을 파악하였다. 본 연구의 종합적인 분석 결과는 다음과 같다.

첫번째, ……(중략)…….

참고 문헌

[1] Mańkowska, M., Pluciński, M., Kotowska, I. and D, L. F. (2021), "Seaports during the COVID-19 pandemic: the terminal operators' tactical responses to disruptions in Maritime supply chains.", *Energies*, 14(14):4339.

[2] Liu, J., Wu, J. J., and Yu, G. (2023), "Maritime supply chain resilience: From concept to practice.", *Computers & Industrial Engineering*, Vol. 182.

[3] Lam, J. and Su, S. (2015), "Disruption risks and mitigation strategies: an analysis of Asian ports.", *Maritime Policy & Management*, Vol. 42, Issue 5, pp. 415-435.

[4] Corrigan, S., Kay, A., Ryan, M., Ward, M. E. and Brazil, B.(2019), "Human factors and safety culture: Challenges and opportunities for the port environment.", *Safety Science*, Vol. 119, pp. 252-265.

…… (중략) ……