

뉴스 데이터 토픽 모델링을 활용한 COVID-19 대유행 전후의 클라우드 보안 동향 파악

저자 : 이선우, 이재우

· 클라우드 보안의 중요성

: 재택 근무 시 직원들이 개인 PC에서 회사 내부 네트워크망을 접속해 생활하는 과정에서 발생할 수 있는 문제 방지.

· 문제

공격자가 회사원의 개인 PC를 통해 사내 네트워크 침입하여 사내 정보 유출, 유출된 정보 기밀이 유출 여부 확인 불가

→ 국내 바이오 기업이 공격 빈도 ↑

· 토픽 모델링 = 비구조화된 문서의 특정 토픽 관련 단어들 강조해주는 것

· 토픽 모델링 방법

① 델파이 기법 (고전적 방법)

> P(문제점) : 절차 복잡, 시간 & 비용 ↑

② LSA (Latent Symantic Analysis)

문서나 단어 행렬 형태로 표현 (DTM : Document for Matrix)

↓

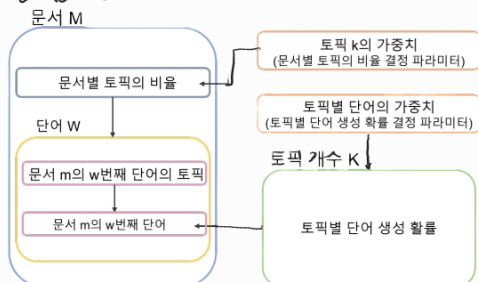
위치는 낮은 가치 갖게 변환

↓

가치 낮은 단어 제거 (TF-IDF)

→ P: 단어 검색 및 처리 속도 ↓, Need to 가중치 부여 & SVD 계산

③ LDA : LSA의 단점을 극복한 토픽 모델링 방법



④ 결과 : 사람들이 COVID-19 후 클라우드를 쓰지 않게 해서 아직 클라우드 보안이 대세 X, 앞으로 중요한 사항이 될 듯.