

기능 명세서

1. 개요

- '기업 뉴스 모니터링 시스템(가칭 RealBrief)'에서 대시보드에 포함시킬 분석 내용을 구현하기 위해 필요한 기능을 명시한 문서임.

2. 기능 명세표

- 아래의 기능 명세표에는 실제 웹 페이지 기능을 제외한 기능을 정리하였음
- 기능 코드와 기능 명시 순서는 실제 워크 플로우를 따르지 않음

* 분석 수행 과정과 상관없이 비슷한 과정을 수행하는 기능끼리 기능 코드를 공유하도록 명시함

기능 코드	기능명	상세	비고
수집-1	뉴스 수집	선택된 기업, 경쟁 기업에 대한 뉴스를 수집합니다.	
전처리-1	데이터 전처리	수집된 뉴스 엔터티의 형식을 일치시킵니다.	
분석준비-1	토픽 모델링/키워드 추출	토픽 모델링과 함께 주요 키워드를 추출합니다.	
분석준비-2	뉴스 감정 태깅 / 감정 점수 할당	뉴스 엔터티들에 대해 감정 분석을 수행하여, 각 엔터티에 감정 태그와 감정 점수(-1, 1, 0 등)를 할당합니다.	
분석준비-3	날짜/키워드/토픽/감정 별 뉴스 클러스터링	필요 시 곧바로 활용할 수 있도록 날짜/키워드/토픽/감정 별로 뉴스 엔터티를 그룹화합니다.	
빈도분석-1	날짜(기간)별 뉴스량, 증감률 계산	날짜(기간)별로 뉴스 수집량, 이전 기간에 비한 증감률을 계산합니다.	
빈도분석-2	부정 키워드/토픽 출현 빈도 계산	사전 정의된 부정 키워드와 토픽이 출현한 빈도(발생량)을 계산합니다.	
빈도분석-3	SoV 지표 계산	SoV 지표 (현재 기업 뉴스량 / (현재기업 + 경쟁사 전체 뉴스량)) 를 계산합니다.	
감정분석-1	평균/가중 감정 점수 계산	매체 영향력 (조회수, 구독수 등)과 조합한 산술 평균을 계산합니다.	매체 영향력 구현방안 강구 필요
감정분석-2	날짜(기간)별 감정 점수, 감정 비율 계산	날짜(기간)별로 감정 점수와 감정 비율을 계산합니다.	
감정분석-3	토픽 별 감정 점수 계산	토픽 모델링을 통해 수행된 토픽 별로 감정 점수를 계산합니다.	
감정분석-4	부정 뉴스 카테고리 분류	감정 분석 결과 부정적 뉴스로 분류된 엔터티들을 사전 정의된 카테고리(법률, 제품, CEO 등)에 따라 분류 합니다.	
감정분석-5	부정 뉴스(키워드) 워드클라우드 생성	감정 분석 결과 부정적 뉴스로 분류된 엔터티들을 활용해 워드 클라우드를 생성합니다.	
트렌드-1	키워드 별 트렌드 수치 계산	각 키워드 별 뉴스 발생량 트렌드 수치를 계산하여 리스트로 저장합니다.	
요약-1	뉴스 요약 파이프라인 구성	뉴스 요약 파이프라인을 구성하여 사용자가 선택한 키워드에 따라 선택된 뉴스 엔터티들을 요약하여 제공할 수 있도록 준비합니다.	