

# 유튜브 스피치 텍스트 변환 툴 기반 인플루언서 전문분야 자동 분류 기법<sup>1)</sup>

구모세(성균관대학교 소프트웨어학과 박사과정)

정재훈(성균관대학교 소프트웨어학과 교수)

## 요약

인플루언서 마케팅의 수요가 높아짐에 발맞춰 인플루언서의 성향을 분석하는 기술이 발전할 필요가 있다. 소셜 네트워크 서비스 기반의 분석은 인플루언서의 포스팅 데이터를 수집하여 분석하는 것으로 신뢰성 있는 분석이 가능했지만 유튜브와 같은 콘텐츠 플랫폼의 데이터는 포스팅 데이터보다 영상의 분석이 중요하다. 본 논문은 인플루언서의 전문성을 분류하기 위한 유튜브 영상 핵심 키워드 수집과 분석 방법을 제시한다.

## 1. 들어가기

인플루언서 마케팅이 중요해지면서, 인플루언서와 클라이언트를 효율적으로 연결하는 매칭 시스템의 필요성이 커지고 있다. 기존 매칭 플랫폼은 NLP와 AI를 활용해 SNS 데이터를 분석, 인플루언서의 성향과 전문성을 자동으로 분류하며, BERT 모델 등을 통해 매칭 정확도를 높였다. 그러나 유튜브와 같은 영상 플랫폼은 텍스트 기반 분석만으로는 전문성을 충분히 파악하기 어려워, 자막(STT) 데이터를 활용한 새로운 분석 기법이 필요하다. 본 논문은 유튜브 자막 데이터를 활용하여 인플루언서의 전문성을 분류하는 새로운 방법을 제안한다. 자막 데이터에서 인플루언서가 반복적으로 언급하는 핵심 키워드를 추출하고, 딥러닝 기반의 BERT 모델[1]을 활용하여 인플루언서의 전문분야를 자동으로 분류하는 프레임워크를 설계하였다. 또한, 자막 데이터에서 불필요한 단어(노이즈)를 제거하기 위해 금칙어 리스트를 적용하여, 의미 있는 키워드만을 필터링하는 전처리 과정을 도입하였다. 이를 통해 유튜브 영상 데이터를 보다 신뢰성 있게 분석하고, 인플루언서의 전문성을 정확하게 파악할 수 있는 새로운 방법론을 제시한다.

## 2. 이론적 배경

SNS 데이터를 활용한 인플루언서와 클라이언트 매칭 시스템은 인스타그램의 빅데이터를 분석하여 인플루언서의 전문분야를 자동 분류하는 기술이다. 이 시스템은 클래스별 인플루언서 데이터를 수집하고 전처리한 후, 라벨링하여 사전 학습된 BERT[1] 모델에 전이학습을 수행해 전문 분야를 자동으로 분류한다. BERT 모델은 Transformer[4] 기반 인코더로, 시계열 데이터의 가중치를 Attention[4] 메커니즘을 통해 학습하며, 1억 개 이상의 파라미터와 대규모 텍스트 데이터를 활용해 문장의 예측을 가능하게 한다. 이를 통해 고도화된 분류 작업이 가능하며, 추가 출력층을 통해 원하는 분류 학습을 수행할 수 있다.

1) 이 논문은 2024년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구임(No. RS-2022-II221199 융합보안대학원(성균관대학교), No. RS-2022-II221015: 지능형 6G 모바일 코어 네트워크 후보 요소기술 개발)

### 3. 연구방법

본 연구의 분석대상은 유튜브에 업로드된 다양한 인플루언서 영상 콘텐츠이다. 특히 본 연구에서는 화장품, 패션, 기술 리뷰와 같은 분야에서 활동하는 인플루언서들의 영상을 주로 선정하였다. 분석대상은 다양한 분야에 걸친 인플루언서들로 구성하여, 각기 다른 주제와 콘텐츠를 분석함으로써 제안된 프레임워크의 일반화 가능성을 검증하고자 한다. 그림1에서 보여지는 것처럼, 자료 수집은 YouTube API를 활용하여 진행되었다. YouTube API를 통해 각 인플루언서의 영상 URL과 해당 영상의 자막 데이터를 수집하였다. 자막 데이터는 구글의 STT(Speech-to-Text) 기술을 통해 자동으로 생성된 서브타이틀을 기반으로 하며, 자막 데이터에는 인플루언서가 영상에서 직접 언급한 주요 내용들이 포함되어 있다. 이는 기존의 제목이나 설명글에 의존한 데이터보다 더 많은 정보를 제공하며, 해당 영상의 주제를 보다 정확하게 분석하는 데 기여한다. 수집된 자막 데이터는 다양한 주제와 분야에 걸쳐 있으며, 각 영상마다 자막 데이터의 양과 구조가 다를 수 있다. 그러나 본 연구는 이러한 자막 데이터를 동일한 프로세스를 통해 전처리하여 분석함으로써 인플루언서의 전문성을 파악하고 분류할 수 있는 방법론을 제시하고자 한다.

본 연구에서는 유튜브 자막 데이터를 기반으로 인플루언서의 전문성을 분류하기 위한 일련의 분석 절차를 진행하였다. 주요 분석 방법은 자막 데이터 전처리, 키워드 추출, 금치어 디셔너리 적용, 그리고 딥러닝 모델 학습 및 결과 분석으로 구성된다.

- (1) 자막 데이터 전처리: 유튜브 API를 통해 수집된 자막 데이터는 KoNLPy 한국어 형태소 분석기를 사용하여 형태소로 분리된다. 자막에서 추출된 데이터는 토큰화 과정에서 품사가 태깅되며, 명사에 해당하는 토큰만을 수집한다. 이때 조사나 감탄사, 불필요한 품사들은 제거하여 데이터의 노이즈를 줄인다. 이를 통해 분석에 필요한 핵심적인 명사 토큰만이 남게 된다.
- (2) 금치어 디셔너리 적용: 수집된 명사 토큰은 금치어 디셔너리 탐색 과정을 거친다. 금치어 디셔너리는 자주 등장하지만 분석에 방해가 되는 단어들(조사, 보조사, 유행어 등)을 포함하며, 이러한 단어들은 모델 학습에서 노이즈로 작용할 수 있다. 금치어 디셔너리에 포함된 토큰은 모두 필터링되며, 금치어에 포함되지 않은 유의미한 키워드만 최종적으로 남게 된다. 그림 2에서 보여주듯, 금치어를 적용하지 않은 경우에는 불필요한 단어들 많이 남아 있으나, 금치어를 적용한 후에는 핵심적인 키워드만 추출된다. 이렇게 추출된 키워드들은 모델의 성능 향상에 기여할 수 있다.
- (3) 딥러닝 모델 학습 및 분류: 추출된 키워드는 사전 학습된 BERT 모델[1]에 입력된다. BERT 모델은 자연어 처리에 강점을 가진 딥러닝 모델로, 추출된 키워드를 기반으로 인플루언서의 전문성을 분류한다. 그림 3에서 보이듯, 기존의 제목과 해시태그만을 기반으로 한 분석 결과는 기타로 분류되는 경우가 많았으나, 본 프레임워크를 통해 추출된 키워드를 사용했을 때는 보다 정확한 분류가 가능했다.
- (4) 성능 비교 및 결과 분석: 실험에서는 제목과 해시태그만을 사용한 기존 방법과, 본 연구에서 제안한 자막 기반 키워드 추출 방법을 비교하였다. 두 가지 방법으로 도출된 분류 결과를

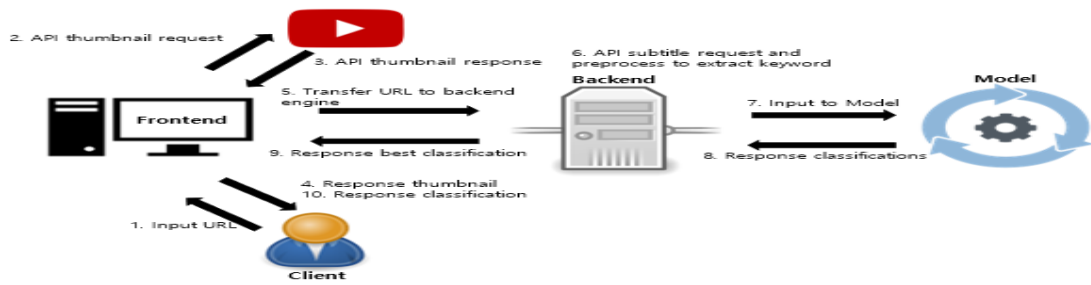


그림 1. 유튜브 자막 데이터 분석 프레임워크

비교한 결과, 본 연구에서 제안한 금치어 딕셔너리와 키워드 추출 과정을 거친 자막 기반 분류 방법이 더 높은 성능을 나타냈다. 이를 통해 본 연구가 제안한 프레임워크의 유효성을 확인할 수 있었다. 이와 같은 절차를 통해 유튜브 자막 데이터를 기반으로 인플루언서의 전문성을 더욱 정확하고 효율적으로 분류할 수 있었으며, 향후 인플루언서-클라이언트 매칭 시스템에서 실질적인 개선을 제공할 수 있음을 실험 결과로 확인하였다.

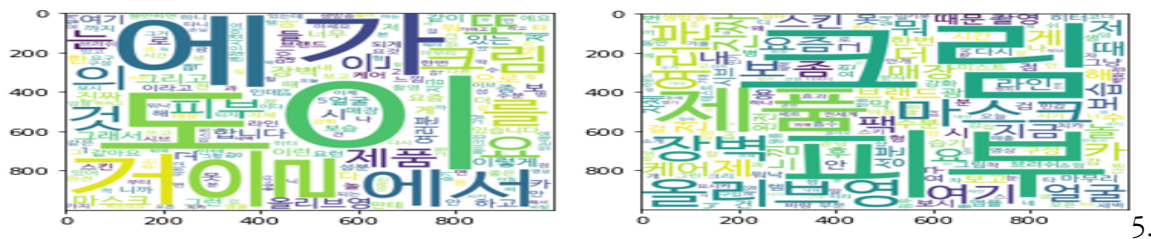


그림 2. 금치어 딕셔너리 성능비교

#### 4. 연구결과

본 연구에서는 유튜브 자막(STT) 데이터를 활용한 인플루언서의 전문성 분류 방법을 제안하고, 금치어 딕셔너리를 적용하여 노이즈를 제거하는 전처리 과정을 통해 성능 향상을 확인하였다. 연구 결과는 다음과 같이 요약할 수 있다. 실험 결과, 금치어 딕셔너리를 적용하여 불필요한 단어(조사, 유행어 등)를 필터링한 후 추출된 핵심 키워드는 인플루언서의 전문성을 분류하는 데 매우 효과적이었다. 금치어를 적용하지 않은 경우에는 불필요한 조사와 단어가 포함되어 있어 분석의 정확도가 떨어졌다. 반면, 금치어 딕셔너리를 적용한 후에는 인플루언서가 실제로 강조한 핵심 키워드만이 추출되었으며, 이를 통해 더 정확한 분류가 가능해졌다. 이러한 핵심 키워드는 딥러닝 모델 입력 시 높은 분류 성능을 보여주었다. 그림 3에서 보이는 실험 결과, 기존의 제목과 해시태그만을 기반으로 한 분석 방법에서는 인플루언서의 전문성을 정확하게 분류하는 데 한계가 있었다. 이로 인해 분류 결과가 "기타"로 출력되는 사례가 많았다. 반면, 자막 데이터를 통해 추출된 키워드를 딥러닝 모델에 입력한 경우, 인플루언서의 실제 전문 분야를 더 정확하게 파악할 수 있었다. 본 연구에서 제안한 자막 데이터를 활용한 분석 방법은 기존의 텍스트 기반 분석보다 더 많은 정보를 제공하며, 인플루언서의 콘텐츠 성격을 더욱 정확하게 파악할 수 있는 방법임을 입증했다. 유튜브 콘텐츠의 경우, 제목이나 설명 글만으로는 인플루언서의 전문성을 파악하기 어렵지만, 자막 데이터를 활용하면 영상에서 실제로 다루는 내용을 분석할 수 있어 보다 깊이 있는 이해가 가능하다. 이는 인플루언서와 클라이언트 간의 매칭 정확도를 높이는 데 기여할 수 있다.