

Sentence-BERT를 활용한 YouTube 가짜뉴스 탐지 시스템 연구

김범중¹, 허지혜, 이협건, 김영운
¹한국폴리텍대학교 서울강서캠퍼스 빅데이터과
 cjawnd7939@gmail.com, floramiss@naver.com,
 luckkim@kopo.ac.kr, hglee67@kopo.ac.kr

A Study On YouTube Fake News Detection System Using Sentence-BERT

Beom Jung Kim¹, Ji Hye Huh, Hyeopgeon Lee, Young Woon Kim
¹Department of Big Data, Seoul Gangseo Campus of Korea Polytechnics College.

요 약

IT 기술의 발달로 인해 뉴스를 제공하는 플랫폼들이 다양해 졌고 최근 해외 인터뷰 영상, 해외 뉴스를 Youtube Shorts 형태로 제작하여 화자의 의도와는 다른 자막을 달며 가짜 뉴스가 생성되는 문제가 대두되고 있다. 이에 본 논문에서는 Sentence-BERT를 활용한 YouTube 가짜 뉴스 탐지 시스템을 제안한다. 제안하는 시스템은 Python 라이브러리를 사용해 유튜브 영상에서 음성과 영상 데이터를 분류하고 분류된 영상 데이터는 EasyOCR을 사용해 자막 데이터를 텍스트로 추출 후 Sentence-BERT를 활용해 문자 유사도를 분석한다. 분석결과 음성 데이터와 영상 자막 데이터가 일치한 경우 일치하지 않은 경우보다 약 62% 더 높은 문장 유사도를 보였다.

1. 서론

IT 기술의 발달로 인해 뉴스를 제공하는 플랫폼들이 다양해졌고 최근 해외 인터뷰 영상, 해외 뉴스를 Youtube Shorts 형태로 제작하여 화자의 의도와는 다른 자막을 달며 가짜 뉴스가 생성되는 문제가 대두되고 있다.

과거 뉴스 제공 수단은 텔레비전, 라디오, 신문을 통해 정보를 제공하고 소비되었지만, 현재는 인터넷 포털 사이트, SNS 메신저를 통해 정보의 제공과 소비가 이루어지고 있다.[1]

가짜 뉴스는 다양한 플랫폼에 의해 정보가 소비됨에 따라, 공신력이 없는 사실과는 다른 뉴스들의 생성과 공유가 쉬워지고 다양해 졌다.

이에 본 논문에서는 Sentence-BERT를 활용한 SNS 가짜뉴스 시스템을 제안한다.

본 논문의 구성은 다음과 같다. 2장에서는 Sentence-BERT의 구성에 대해 살펴보고, 3장에서는 Sentence-BERT를 활용해 문장 유사도 분석 시스템을 제안한다. 마지막 4장에서는 결론 및 향후

연구 과제를 제시한다.

2. 관련연구

본 장에서는 Sentence-BERT의 Cross-Encoder와 Bi-Encoder에 대해 알아본다.

2.1 Cross-Encoder

Cross-Encoder[2]는 두 문장을 Tranformer 네트워크에 전달한다. 두 개의 문장은 모델에 넣어 문장이 변형되지 않는 상태에서 유사도 비교를 진행한다. 정보 손실 없이 문장간 관계를 파악하는데 성능이 우수하나 비교해야 하는 문장이 증가하면 연산이 급증한다.

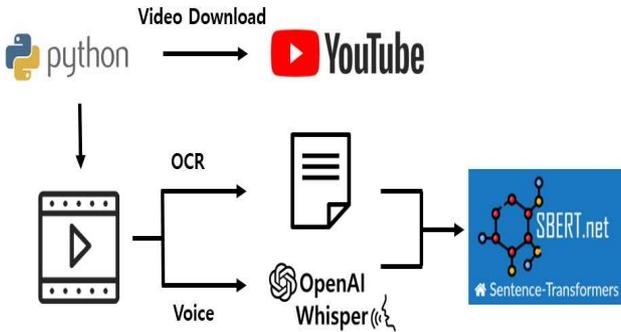
2.2 Bi-Encoder

Bi-Encoder[2]는 두 문장에 대해 각각 별도의 인코더 모델을 사용한다. 따라서 두 문장의 다른 벡터 값이 생성되고 Cosine 유사도 분석을 통해 해당 문장의 유사도를 계산한다.

3 제안하는 가짜뉴스 탐지 시스템

문장 유사도 검사는 YouTube로부터 영상을 다운로드 받아 음성 데이터와 영상 자막데이터로 분류 후 OCR과 음성 번역을 통해 텍스트로 변환하여 Sentence-BERT를 통해 문장 유사도 검사를 진행한다.

[그림 1] 제안하는 가짜뉴스 탐지 시스템을 나타낸다.



[그림 1] 제안하는 가짜뉴스 탐지 시스템

뉴스 데이터 수집은 파이썬 pytube 라이브러리를 사용해 YouTube 동영상 다운로드 후 영상과 음성 데이터를 분리한다.

음성 데이터는 python openai 라이브러리를 사용해 OpenAI Wishper로 전송해 한국어로 번역된 결과를 가져온다.

영상 데이터는 자막이 송출되는 영역에 ROI를 설정하여 특정 영역을 분리한 후 Image로 저장 후 EasyOCR을 통해 해당 영역의 텍스트를 분리한다.

문장 유사도 분석은 Sentence-BERT의 Bi-Encoder 모델을 사용해 두 문장의 Cosine 유사도를 분석한다.

3.1 문장 유사도 분석

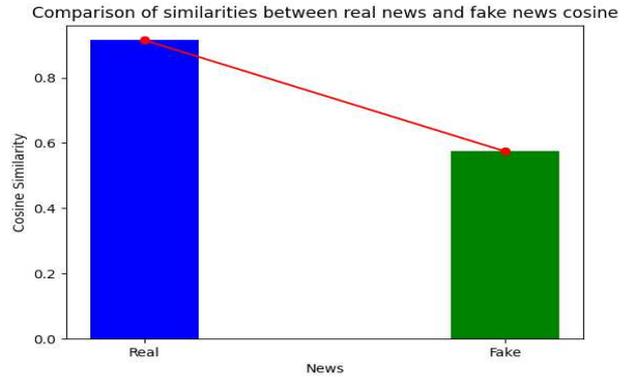
문장 유사도 분석은 실제 뉴스 데이터와 가짜 뉴스 데이터를 분석하였다.

뉴스 데이터 수집은 [그림 1]을 기준으로 데이터를 수집하였다.

다음<표 1>과 [그림 2]는 실제 뉴스 데이터와 가짜 뉴스 데이터의 문장 유사도 분석 결과이다.

<표 1> 문장 유사도 분석 결과

항목	Cosine Similarity
Real News	0.915404200553894
Fake News	0.5741439461708069



[그림 2] 문장 유사도 분석 결과

문장 유사도 분석 결과 음성 데이터와 영상 자막 데이터가 동일한 경우 일치하지 않는 경우보다 약 62% 더 높은 유사도를 보였으며, 음성 데이터와 영상 자막 데이터가 더 낮은 경우 매우 낮은 유사도를 보였다.

4. 결론 및 향후 연구과제

본 논문에서는 Sentence-BERT를 활용한 YouTube 가짜 뉴스 탐지 시스템을 제안한다.

Python 라이브러리를 사용해 YouTube 영상에서 영상 데이터와 음성 데이터를 분류 후 OCR과 Wishper를 사용해 실제 뉴스 데이터와 가짜 뉴스 데이터를 문장으로 비교한 결과 음성 데이터와 영상 자막 데이터가 일치한 경우 일치하지 않은 경우 보다 62% 높은 유사도를 보였다.

하지만 본 논문의 음성 데이터는 다국적 언어를 한국어로 번역 후 유사도를 비교하였기 때문에 향후, 다양한 뉴스 영상 데이터 세트를 통해 시스템에 대한 성능 평가가 필요하다.

참고문헌

[1] 남유원, 지성욱. 뉴스 이용, 참여, 출처 인지가 뉴스 미디어 신뢰도에 미치는 영향 : SNS, 메신저, 유튜브를 통한 뉴스 이용을 중심으로. 한국언론학보, 66(5), 57-94. 2022

[2] Reimers, Nils, and Iryna Gurevych. "Sentence-bert: Sentence embeddings using siamese bert-networks." arXiv preprint arXiv:1908.10084 (2019).