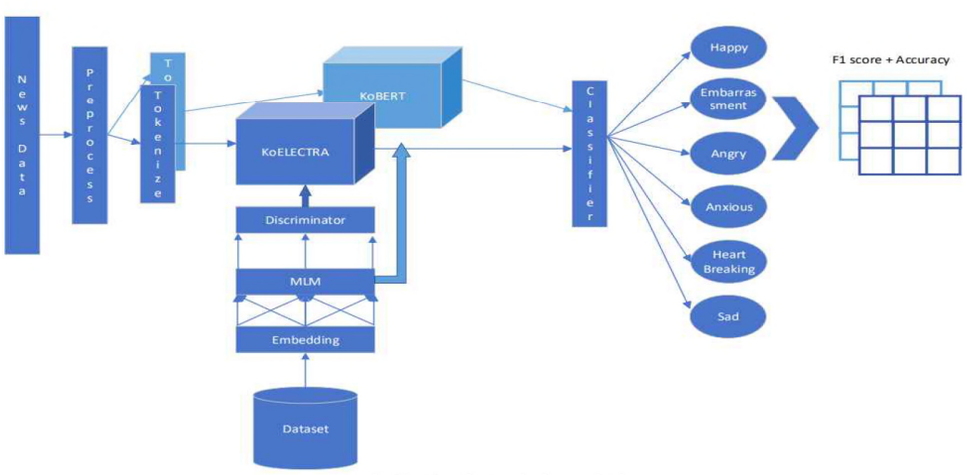
딥러닝 기반 메타버스 뉴스 텍스트 감정 분석[1)]

차현주[2)](성균관대학교 미디어문화콘텐츠연구소 겸임교수)
권상희[3)](성균관대학교 미디어커뮤니케이션학과 교수)
구모세, 권준희(성균관대학교 소프트웨어학과), 정재훈(성균관대학교 소프트웨어학과 교수)

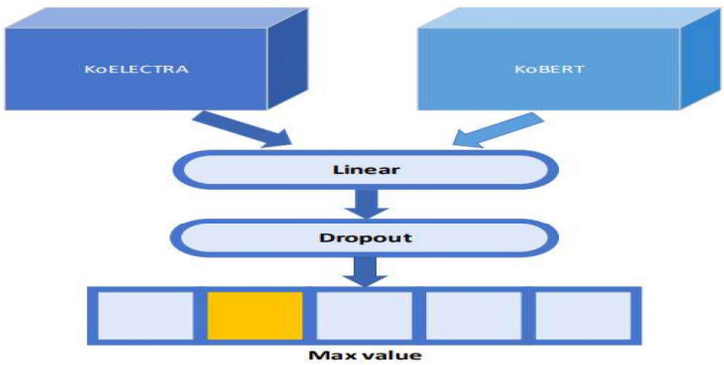
요 약

인공지능과 소셜미디어 발달로 인해 방대한 양의 데이터가 축적되는 뉴스에서는 사람들의 의견과 태도, 성향 등과 같은 주관적인 데이터로부터 객관적인 형태의 정량화된 정보를 분석하기 위한 기술로 감정분석(Emotional Analysis)의 비중이 부각되고 있다. 극성(Polarity)을 판단하여 긍정과 부정 또는 중립으로 분류하는 비교적 단순한 감성분석(Sentiment Analysis)과는 달리, 감정분석(Emotional Analysis)은 텍스트에서 느껴지는 감정을 보다 다양하고 정교하게 분류된 정보를 제공하고 이를 유용하게 활용할 수 있다는 점에서 매우 중요하다. 더욱이 최근 자연어처리(Natural Language Processing; NLP)의 발달로 인공지능(AI)은 길고 복잡한 문장의 맥락을 이해하고 의미를 해석하는 수준에 이르렀다. 특히, 트랜스포머 계열인 KoBERT와 KoELECTRA는 학습데이터로 사전 훈련된 AI 모델은 약간의 미세조정(fine-tuning)을 통해 높은 분류 성능을 낼 수 있기 때문에 다중 감정 분류에 적합한 모델이다. 그러나 기존 연구들은 대체적으로 각 문서 최소 단위인 어휘의 감성 극성에 기반한 감성분석(Sentiment Analysis)에 주목하였으나 본 연구는 뉴스 텍스트에서 느껴지는 여러 가지 감정을 세부적으로 판별할 수 있는 감정분석(Emotional Analysis)에 초점을 두고 딥러닝 기반 메타버스 뉴스 텍스트 감정분석 모델을 개발하였다는 점에서 다른 연구들과 차별화된다. 이에 본 연구는 메타버스 뉴스 데이터를 대상으로 6개 기본 감정(기쁨, 당황, 분노, 불안, 상처, 슬픔)을 기준으로 각각 9개의 세부 감정이 포함된 총 60개 감정 상태를 분류하는 딥러닝 기반 감정분석 모델을 구축하고, 감정분석 성능을 높이기 위해 KoBERT와 KoELECTRA 모델간의 성능 비교를 통해 견고하고 좋은 성능을 낼 수 있는 감정 분석기법을 제안한다.



[그림 1] 본 연구 파이프 라인

1) 이 논문은 성균관대학교의 2022학년도 AI융합연구비에 의하여 연구되었음
2) top1@skku.edu
3) skweon@skku.edu



[그림 2] KoBERT와 KoELECTR

<표 1> 본 연구 감정 분류

60가지 감정 분류						
기분	기쁨	당황	분노	불안	상처	슬픔
1	감사하는	고립된	툴툴대는	두려운	질투하는	실망한
2	사랑하는	남의 시선 의식하는	좌절한	스트레스 받는	배신당한	비통한
3	편안한	외로운	짜증나는	취약한	고립된	후회되는
4	만족스러운	열등감	방어적인	혼란스러운	충격 받은	우울한
5	흥분되는	죄책감	악의적인	당혹스러운	불우한	마비된
6	느긋한	부끄러운	안달하는	회의적인	희생된	염세적인
7	안도하는	혐오스러운	구역질나는	걱정스러운	억울한	눈물이 나는
8	신이 난	한심한	노여워하는	조심스러운	괴로워하는	낙담한
9	자신하는	혼란스러운	성가신	초조한	버려진	환멸을 느끼는