자연어 처리 기술을 이용한 불공정 약관 판별 서비스

박단비, 최민서, 정성호
한국외국어대학교[1]

sugarrain00@gmail.com, minmeanseo@gmail.com, shjeong@hufs.ac.kr

Unfair Terms and Conditions Identification Service Using Natural Language Processing Technology

Dan Bi Park, Min Seo Choi, Seong Ho Jeong
Hankuk University of Foreign Studies[1]

요 약

최근 들어 소비자들의 불공정 약관 계약 체결에 따른 피해 사례가 증가하고 있다. 본 논문에서는 소비자의 불공정 약관 체결 사례를 사전에 예방하고자 자연어 처리 기술을 활용한 불공정성 판별 모델 및 서비스 구조를 제시한다. 이를 위해 한국어 데이터 특화 처리를 위하여 KoBERT 모델을 사용하고, 서비스 제공 과정에서 공정거래위원회의 불공정 약관 데이터를 전처리 및 학습한다. 아울러 사용자 편의성 증대를 위하여 구현된 웹페이지 및 크롬 확장 프로그램도 제시한다.

Ⅰ. 서론

본 논문에서는 소비자들의 불공정 약관 계약 체결에 따른 피해 사례가 증가함에 따라, 자연어 처리 기술을 활용하여 이를 사전에 판별하고 예방할 수 있는 서비스 구조를 제시한다. 본 연구는 약관의 불공정성을 자동으로 판별할 수 있는 이진 분류모델을 학습시켜 활용하는 것을 목표로 하며, 불공정 약관 판별 모델 연구 및 웹 서비스 개발 2 가지 단계로 나누어 진행하였다. 서비스 제공 과정에서 판별 모델의 정확도 향상을 위하여 공정거래위원회에서 제공하는 불공정 약관 조항에 대한 의결서와 시정 보도자료를 통해 데이터를 확보하였다. 대용량의 텍스트 형태로 존재하는 약관 데이터를 이용하기 위해 자연어 입력을 특정 확률로 반환하는 함수인 자연어 처리 모델인 KoBERT 를 활용하였다.

Ⅱ. 불공정 약관 판별 모델

모델을 개발하기 위해 수집한 데이터셋의 데이터 불균형 및 부족 문제를 해소하고 모델의 판별 성능을 향상시키기 위해 데이터 증강, 사용자 사전 정의, 불용어 제거 기법을 통해 데이터의 전처리를 수행한다. 데이터 증강 기법으로는 EMNLP 에서 발표한 EDA 중 RS(Random Swap)와 RD(Random Delete) 기법을 활용한다[1]. KoNLPy 의 Mecab-Ko 형태소 분석기를 통해 사용자 사전 정의 및 불용어 제거를 진행하며 모델의 판별 성능을 향상시킨다. 전처리한 데이터를 기반으로 한국어 특화 BERT 모델인 KoBERT(Korean BERT pre-trained cased)의 이진 분류 지도학습을 실시한다. 단순히 이진 분류를 하는 것이 아닌, 입력 데이터의 불공정성 판별을 백분율로 명확히 표기하기 위해서 Sigmoid 활성 함수를 활용하였고, 학습된 모델은 최종적으로 93.18%의 정확도를 보였다.

Ⅲ. Chrome Extension 을 활용한 웹 서비스 개발

최종 학습 모델은 h5 파일로 저장된다. Python 웹 프레임워크인 Flask 로 저장 모델을 serving 하였고, 웹 및 google chrome extension 으로 불공정 약관 판별 서비스를 활용할 수 있도록 구현하였다.

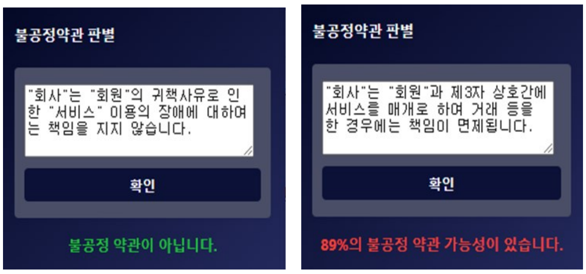

그림 1. Chrome extension 약관 판별 구현 예시

Ⅳ. 결론

본 논문에서는 KoBERT 모델 및 자연어 처리 기법 활용을 통하여 불공정 약관 판별 모델 및 서비스를 개발하였다. 이를 통해 소비자 스스로 약관의 불공정성을 판별할 수 있는 토대를 마련하였다.

참 고 문 헌

[1] Jason Wei·Kai Zou. (2019). EDA : Easy Data Augmentation Techniques for Boosting Performance on Text Classification Tasks, EMNLP-IJCNLP 2019 short paper.