

소웨지존

2019년 1학기 인공지능(001): 팀 프로젝트
Github 주소: <https://github.com/hyj378/ArtificialIntelligence>

[1팀] 이모티콘 추천 분류기

2019년 4월 29일 ~ 2019년 6월 12일

1. 프로젝트 개요

- 프로젝트 소개(주요내용)

이모티콘을 주석으로 달아놓은 상태에서 문장을 미리 설정해 놓은 클래스로 분류하고, 클래스에 맞는 이모티콘을 추천하는 AI분류기 제작 프로젝트

- 프로젝트 제안배경(주제 선정의 이유 / 목표)

-이모티콘을 많이 사용하게 되고, 캐릭터의 인기도 높아지면서 이모티콘은 문자에 필수가 되었다. 그러나 유용한 기능인 이모티콘 추천기는 현재 한국어로 지원되지 않는다. 이제는 한국어문장의 의미를 분류하는 모델을 만들어보고, 나아가 이모티콘을 추천해보자

2. 주요기능 및 예상 결과물

- 주요 기능 : 메신저 어플리케이션 사용자가 메세지 문장을 작성했을 시, 그 문장을 분석하여 메세지의 감정/분위기에 적절한 이모티콘들을 사용자에게 추천해준다.
- 예상 결과물
 1. 문장의 분위기에 맞는 적절한 이모티콘을 사용자에게 보여주고, 메세지와 함께 보낼 수 있도록 한다.
 2. 문장의 분위기에 맞는 적절한 이모티콘을 사용자가 수동적으로 검색할 수 있게 하고, 검색 시에 이모티콘들을 찾아준다.

3. 기대효과 및 활용분야

- 기대효과

- 사용자가 전하고자 하는 말을 분석하여 사용자의 감정에 맞는 이모티콘을 추천해주기 때문에 사용자가 직접 이모티콘을 찾지 않아도 되는 편리함이 있다.
- 사용자의 감정을 이모티콘을 통해서 효과적이고, 빠르게 전달할 수 있다.
- 사용자들의 이모티콘 사용 빈도수가 늘어남으로써 전보다 더 이모티콘을 찾게 되고, 이모티콘 생산자들도 그에 맞춰 여러가지 감정을 갖는 이모티콘을 만들어 내면서 생산성이 증대한다.

- 활용분야: 카카오톡, 인스타그램, 페이스북 등 이모티콘이 사용되며 소통이 활발하게 일어나는 곳에서 메시지 감정 분석을 통한 이모티콘 추천 AI분류기를 적용할 수 있다.

4. 역할분담

- 김경남 (소프트웨어학과 17011654)
자료 조사, 모델 제작, UI 제작
- 김민주 (소프트웨어학과 17011660)
자료 조사, 모델 제작, UI 제작
- 유지인 (디지털콘텐츠학과 16011191)
자료 조사, 데이터 가공, UI 제작
- 이주혁 (소프트웨어학과 17011683)
자료 조사, 데이터 가공, UI 제작
- 황유진 (지능기전공학부 무인이동체공학 18011756)
자료 조사, 모델 제작, UI 제작

5. 프로젝트에 사용할 데이터 셋

- 트위터 데이터 이용 - 트위터 API
- 한국어 감정 분석 말뭉치 - <http://word.snu.ac.kr/kosac/>
- 국립국어원 언어정보나눔터 - <https://ithub.korean.go.kr/user/main.do>
- 한국어 감성 사전 - <http://dilab.kunsan.ac.kr/knusl.html>
- 형태소 분석 및 품사 태깅 제공(데이터전처리) - <https://konlpy-ko.readthedocs.io/ko/v0.4.3/morph/>

6. 기존 연구 / 개발 조사 결과

• 기존연구

MOJITALK: Generating Emotional Responses at Scale 저자 [Xianda Zhou](#), [William Yang Wang](#)

기존의 emotional language는 데이터가 부족했고, 라벨도 한정적("happiness", "sadness", "anger" 같이 단순하거나 아예 "positive", "negative"처럼 이종분류)이었다. 이 연구에서는 다양한 사람들이 다양한 낱말로 직접 이모티콘을 붙인 트위터 데이터를 전처리해 training data 로 사용했고, 주어진 응답으로부터 emotion(이모티콘)을 생성하기 위해 강화된 CVAE모델(조건부 VAE)과 같은 최첨단 신경 모델을 이용했다.

• 조사결과

우리 조가 직면했던 문제(이미 존재하는 데이터들의 라벨링의 한계)를 해결하기 위해 트위터 데이터셋을 이용하면 좋을 것 같다.

이모티콘 생성을 위해 CVAE모델(또는 VAE)을 이용을 고려하자