

2020 국어 정보 처리 시스템 경진 대회 보고서

TEAM 감정저금통

서론

코로나로 인해 사회활동이 줄어들고, 그로 인한 여파로 사람들의 우울지수가 늘어나는 일명 코로나 블루가 사회적인 이슈입니다. 저희는 이에 조금이나마 도움이 될 수 있도록 국민들이 가장 많이 이용하는 메신저인 카카오톡을 이용한 감성분석 일기 챗봇을 만들었습니다.

챗봇으로 일기를 입력하면, 감성분석 AI가 긍정/부정 여부를 판단합니다. 부정인 경우, 좋은 글귀와 본인의 저금용 계좌로 송금할 수 있는 버튼을 보내줍니다. 설정한 목표금액에 도달하면 알려주어서 목표달성의 성취감과 기쁨을 느끼며, 누적된 스트레스와 우울감을 해소할 수 있다고 생각합니다. 또한 매주 일요일마다 한주동안 쓴 일기의 분석결과를 바탕으로 그래프도 그려줍니다. 이를 통해 본인의 일주일을 돌아볼 수 있습니다.

인공지능

카카오톡 채널 "감성 분석 일기장"에서 인공지능 분석을 담당하고 있는 서버입니다.

인공지능은 ResNet과 CNN을 이용한 신경망 1과, Resnet과 ELMO 개념을 응용한 신경망 2가 각자 값을 산출한 뒤, 두 신경망의 평균을 출력하게끔 설계했습니다.

데이터

데이터는 Naver sentimental Movie corpus v1.0 의 ratings_train.txt, ratings_test.txt를 사용했습니다.,

● 전처리

1. 중복되는 댓글 및 비어있는 항목 제거
2. 한글로 된 댓글로만 학습하기 위해 정규표현식으로 한글이 아닌 항목 빈 값으로 지정

대한민국 영화 수준의 평균치를 하락시키기에 충분한 c급 영상물
2009년에 만들어진것치고 재밌음 영화가 길어서 좀 다들었으면 ..소재도 좋고 몰입감도 기대이상 매끄럽지 못한 부분도 있지만 연기는 둘 다 잘한듯
대한민국 영화 수준의 평균치를 하락시키기에 충분한 급 영상물
년에 만들어진것치고 재밌음 영화가 길어서 좀 다들었으면 소재도 좋고 몰입감도 기대이상 매끄럽지 못한 부분도 있지만 연기는 둘 다 잘한듯

3. '<s>'라는 start token 과 '<e>'라는 end token 추가
4. 한국어는 오타, 신조어, 줄임 말도 있으므로 단어 단위가 아닌 음절 단위로 처리

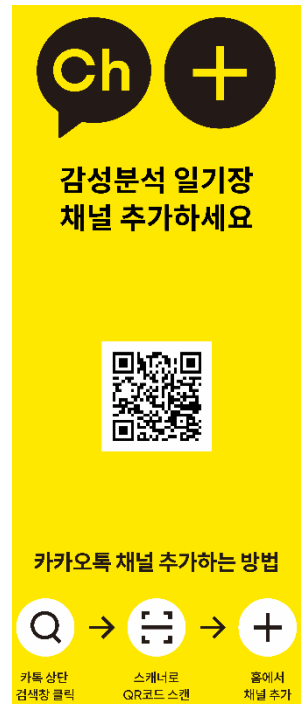
['<s>', '아', ' ', '더', '빙', ' ', '진', '짜', ' ', '짜', '증', '나', '네', '요', ' ', '목', '소', '리', '<e>']

5. Tokenizer 를 이용해 단어를 숫자로 변경

[3, 15, 2, 90, 495, 2, 53, 81, 2, 81, 301, 17, 52, 32, 2, 310, 99, 26, 4]

6. 패딩을 통해 길이 102 으로 설정

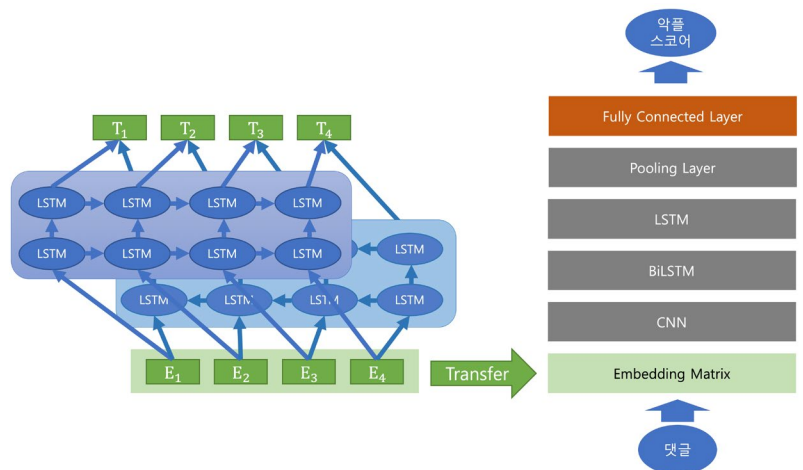
```
[ 3 15 2 90 495 2 53 81 2 81 301 17 52 32 2 310 99 26
 4 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0]
```



■ 신경망에 이용한 개념

● ELMO Architecture

ELMO 는 그림과 같이 Bi-LSTM 을 활용한 다층의 LSTM 의 결과를 취합한 모델입니다.

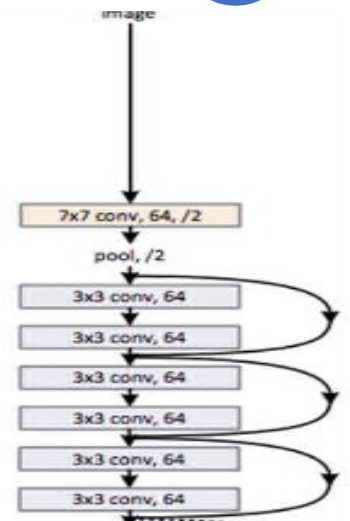


● ResNet Architecture

ResNet 은 ILSVRC 2015 대회에서 우승한 네트워크 방식입니다.

ResNet 은 skip connection 을 이용해 잔류 표현 함수를 학습함으로써, 매우 깊은 신경망을 구성 할 수 있습니다.

이 구조를 이용해서 얻는 이득으로, layer 가 깊을 때 일어나는 gradient vanishing 문제를 어느정도 방지할 수 있습니다.



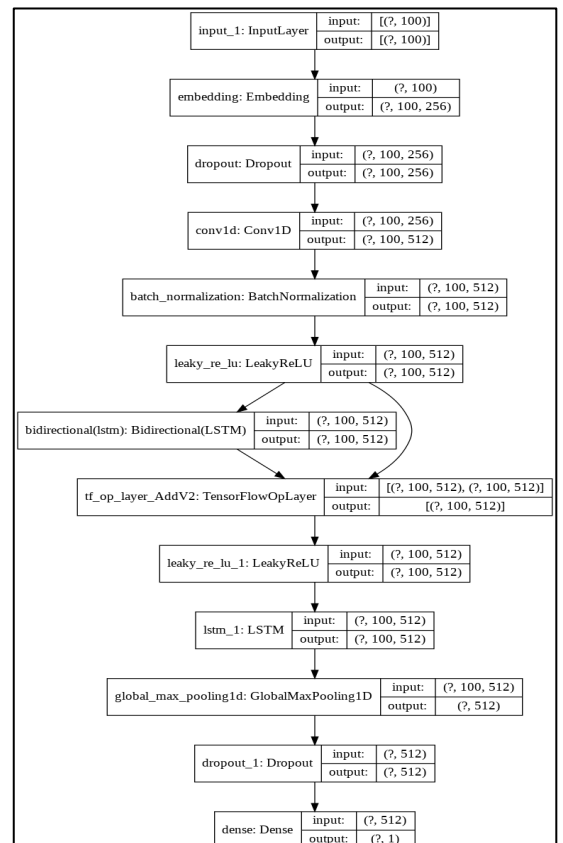
[신경망 - 1] (ResNet 과 ELMO 개념을 이용)

1. 들어온 문장을 Embedding Layer 를 통해 벡터 화 시킨 후 Convolution Layer 로 특징을 잡아줍니다.
이후 batch normalization Layer 를 통해 값을 regularization 시켜줍니다.

2. skip connection 을 이용해 Bi-LSTM Layer 를 통해 문장의 앞뒤 문맥을 학습합니다.
이후 해당 layer 를 거친 데이터와 거치지 않은 데이터를 합치고 LSTM Layer 를 한번 더 거칩니다.

3. Max pooling 을 통해 다시 한번 해당 데이터들의 특징을 추출한 뒤 Full connection layer 를 통해 학습해줍니다.

해당 모델은 약 86% 의 acc 를 보인다 (accuracy 0.866)



[신경망 - 2] (ResNet 과 CNN 개념을 이용)

1.Embedding Layer 를 이용해 Tokenizing 된 데이터를 벡터로 변환

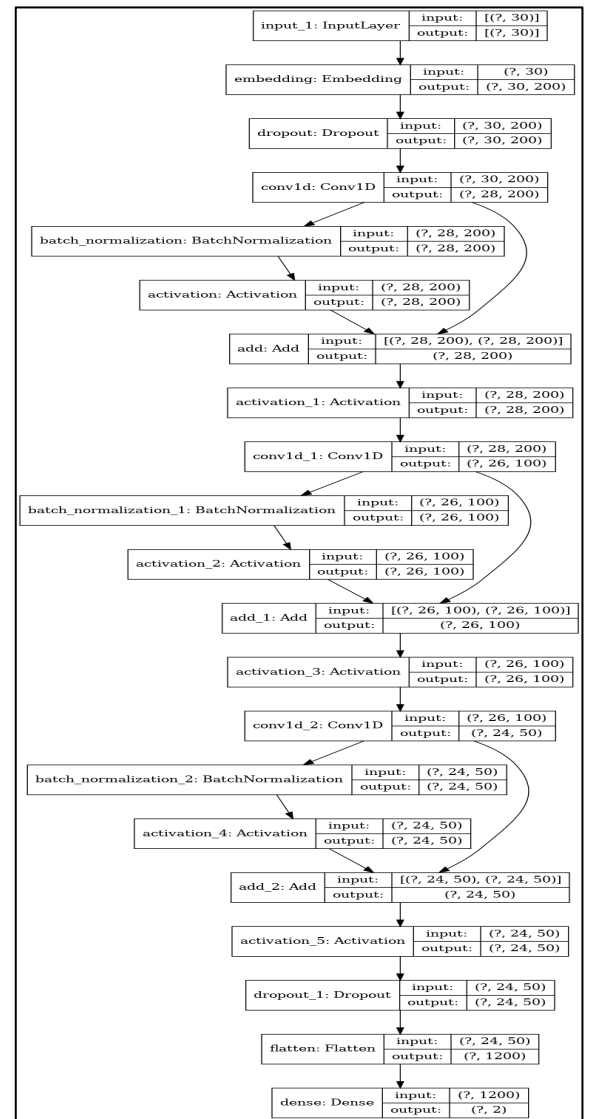
2.이후 ResNET 구조를 이용해 Convolutional Layer 의 학습 전/후 값을 더하는 block 을 3 개 둔다.

3.모든 block 의 전/후에 Dropout 을 이용해 overfitting 을 방지한다.

4.sigmoid 를 이용해 최종 노드의 출력 값을 0 과 1 사이로 고정시켰으며, 0 이면 부정, 1 이면 긍정을 출력하게끔 되어 있다.

각각의 Conv layer 에서 학습된 데이터와 학습하지 않은 데이터를 합침

약 85%의 acc 를 보인다 (accuracy 0.854)



[서버 – Flask 이용]

<http://193.122.122.61/gamsung> 에서 서비스하며,

POST 방식의 요청에만 동작하도록 설계했다.

POST 방식으로 { "text" : 감성분석_내용_텍스트 } 형식으로 보내면,

{ "result" : "긍정/부정 결과값" } 으로 주며, 결과값은 긍정일 경우 1, 부정일 경우 -1 을 return 한다.

+ 만약 GET 으로 요청하면, Wrong access 를 return 한다.

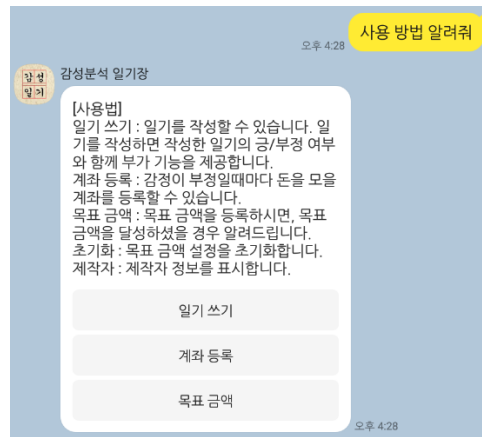
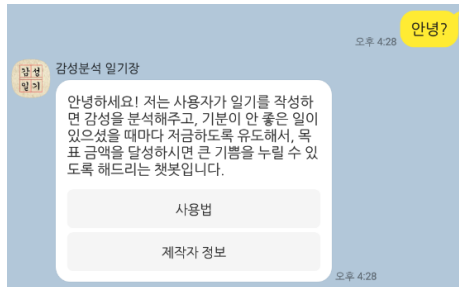
→ 메인 챗봇 서버에서 질의를 통해 감성의 return 값을 전달한다.

KakaoTalk ChatBot

카카오톡 채널 "감성분석 일기장"에서 서비스되고 있는 챗봇입니다.

기본적으로 인사말, 사용 방법, 제작자 정보를 알려주는 기능이 있습니다.

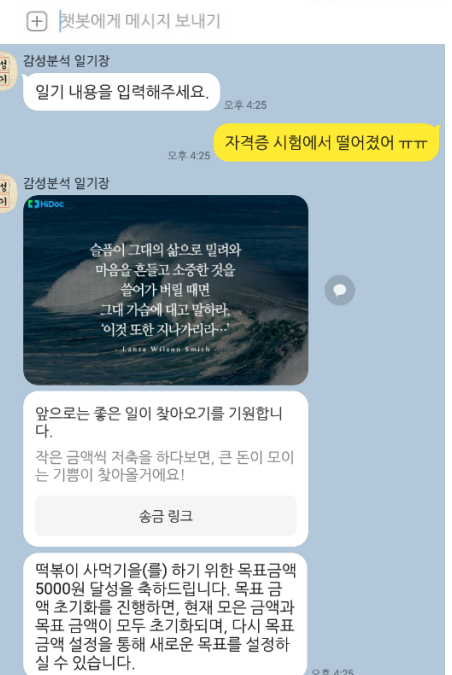
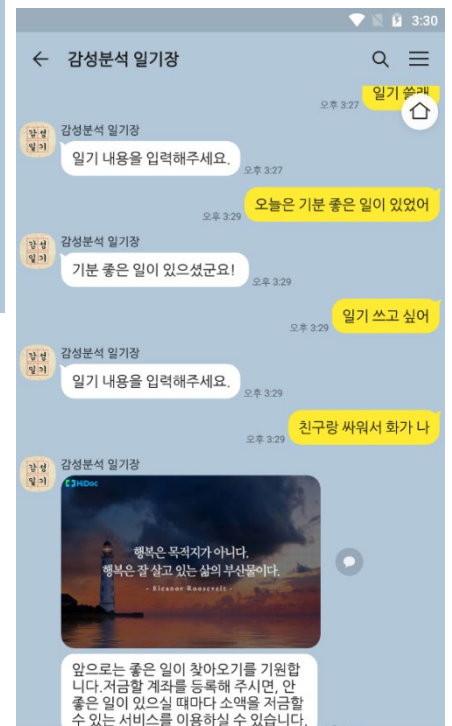
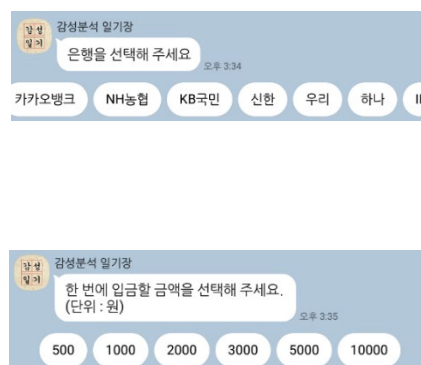
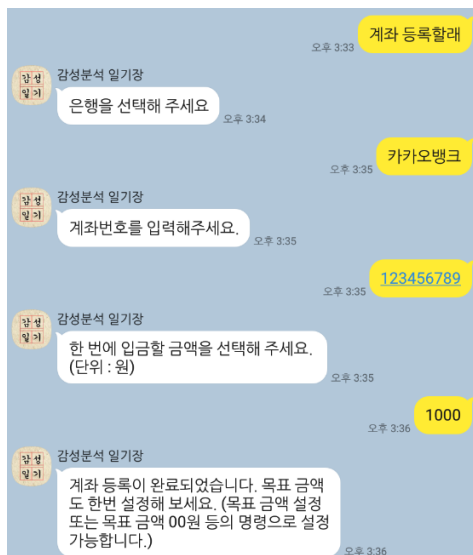
핵심 기능으로는 일기를 작성하면 감성을 분석하여 긍정/부정 여부를 알려주며, 부정일 경우에는 부가적인 기능으로 좋은 글귀가 적힌 그림을 보여주고 사전에 설정한 금액을 저금할 수 있도록 유도합니다.



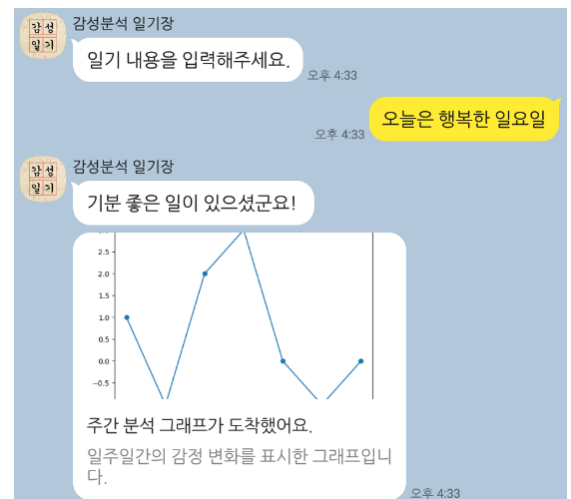
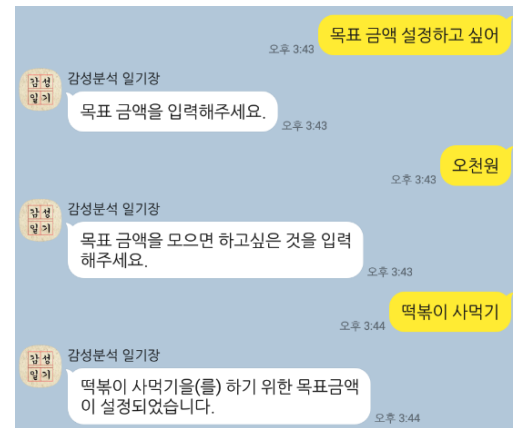
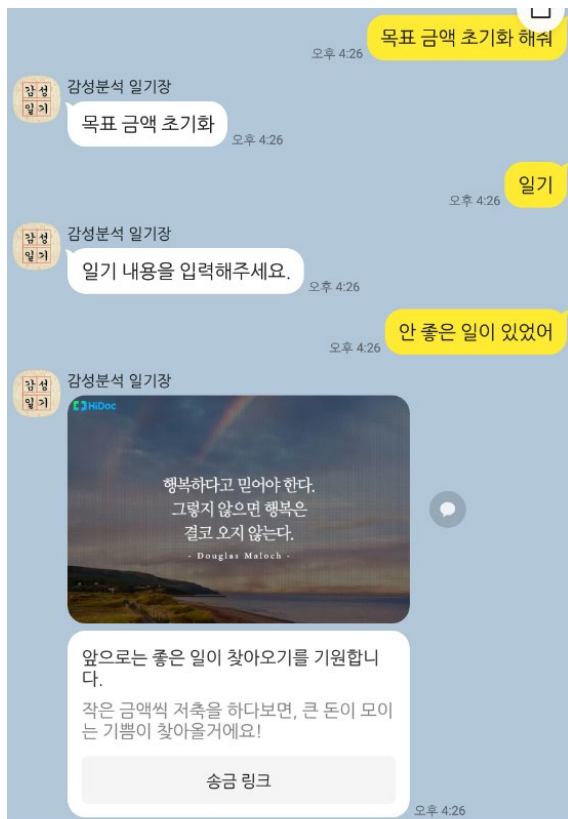
계좌 등록 기능에서는 사용자가 은행, 계좌번호, 한 번에 저금(입금)할 금액을 입력합니다.

은행과 금액은 목록에서 선택할 수 있습니다.

계좌 등록을 하면, 일기 내용이 부정일 경우 좋은 글귀가 적힌 그림과 함께 저금을 할 수 있는 버튼이 제공됩니다. 저금을 할 수 있는 버튼은 사용자가 입력한 정보를 바탕으로 모바일 송금에서 많은 사람들이 이용하는 APP 인 TOSS 의 송금 링크가 제공됩니다. 따라서 TOSS APP 이 설치되어 있는 상태에서 송금 링크 버튼을 터치하기만 하면, 바로 지정된 계좌로 지정된 금액을 송금할 수 있는 화면이 나옵니다.



목표 금액 설정 기능을 이용하여, 목표 금액과 목표 달성 시 하고싶은 것을 입력하면, 목표 금액이 모였을 경우 알려줍니다. 목표 금액 달성 여부는 사용자에게 송금 링크 버튼을 전송했을 때, 모두 송금을 완료했다고 가정합니다.



매 주 일요일마다 첫 번째 일기를 작성했을 때, 일주일간의 감정 변화 그래프를 보여줍니다.

☞ 챗봇에게 메시지 보내기

목표 금액 초기화 기능을 이용하여, 설정된 목표 금액을 초기화 할 수 있습니다. 목표 달성 이후에 사용하면, 더 이상 목표 달성을 축하한다는 메시지가 오지 않게 됩니다. 이 기능을 사용하면 현재까지 모은 금액도 초기화되므로 다시 목표 금액 설정 기능을 이용해서 새 목표를 설정할 수 있습니다.

DB Server

Chatbot Server 와 AI Server 사이에서 데이터 처리 및 저장 역할을 수행한다.

1. 챗봇 사용자의 기본 정보를 저장
2. 계좌, 목표 등 부가 정보를 저장
3. 챗봇 서버에서 일기 내용을 전송 받으면, AI서버로 보내서 긍정/부정 여부를 전송 받아서 저장
4. DB에 있는 정보를 바탕으로 챗봇 응답에 필요한 정보를 챗봇 서버로 전송

