

너목들 개선/보완점

너목들

01. 개선/보완점

NLP

C	D	E	F	G	H	I	J
연령	성별	상황키워드	신체질환	감정_대분류	감정_소분류	사람문장1	시스템응답1
청년	여성	진로, 취업, 직장	해당없음	기쁨	신이 난	지금 난 기분이 너무 좋아.	굉장히 즐거우신 것 같은데요? 좋은 일이 있나요?
청년	여성	연애, 결혼, 출산	해당없음	기쁨	신이 난	나도 조카가 생겨! 너무 기뻐.	정말 좋은 소식이에요. 축하할 특별한 계획이 있을까요?
중년	여성	직장, 업무 스트레스	해당없음	불안	불안	거래처와의 다음 계약이 무산될까봐 불안해.	많이 걱정되시겠어요. 계약이 무산될 것 같다고 생각하는 이유가 있으신가요?
노년	남성	대인관계	만성질환 유	슬픔	엄세적인	당뇨로 고생 중인데 아내가 이혼하자더군. 아내가 날 버리면 병은 더 악화될 텐데 세상에 나 혼자 남은 기분이야.	몸이 아픈데 아내가 이혼하자고 하니 혼자 있는 것 같은 기분을 느끼셨군요.
노년	남성	재정	만성질환 유	슬픔	마비된	노년에 재정적으로 이루어 놓은게 없어서 걱정이야.	나이들어 재정적으로 어려워서 걱정되시는군요.
노년	여성	대인관계	만성질환 유	불안	조심스러운	요즘 코로나 때문에 김 할아버지를 만나러 가지 못하고 있어.	그렇군요. 많이 슬프시겠어요.
청년	남성	대인관계	해당없음	당황	혼란스러운	하루 만에 부서 동료가 모두 다른 사람으로 바뀌어서 당황스러웠어.	부서 동료가 모두 바뀌었다니 정말 당황했겠어요.
청소년	여성	가족관계	해당없음	상처	괴로워하는	오늘 오빠가 나 때문에 얼마랑 싸웠어.	정말요? 많이 속상하셨겠어요.
청년	남성	대인관계	해당없음	슬픔	슬픔	최근 물난리로 인해 직장동료가 죽었어. 다들 슬픔에 잠겨 있어.	직장동료분이 돌아가셔서 주위가 슬픔에 잠겨있군요.
노년	남성	건강, 죽음	만성질환 유	분노	구역질 나는	아내가 검사를 받고 싶다고 했는데 약국에서 약만 사주다가 결국 큰 병으로 만들고 말았어.	아내에게 큰 병이 생겨 마음이 괴로우시겠어요.
중년	여성	직장, 업무 스트레스	해당없음	분노	악의적인	박 과장님이 나에게 너무 모욕적인 말을 했어.	앞에서 모욕적인 말을 들어서 화나셨겠어요.

<감성 대화 말뭉치 데이터셋>

감성 대화 말뭉치 데이터셋의 경우 위와 같이 구성되어 있습니다.
 하지만, 고민이 아닌 문장에 대해서는 챗봇이 엉뚱한 대답을 할 수 있어서
 일상 대화에도 자연스럽게 응답할 수 있도록 모두의 말뭉치에서 제공하는 **일상 대화 말뭉치**도 추가하여 학습하려고 합니다.

01. 개선/보완점

NLP

C	D	E	F	G	H	I	J
연령	성별	상황키워드	신체질환	감정_대분류	감정_소분류	사람문장1	시스템응답1
청년	여성	진로, 취업, 직장	해당없음	기쁨	신이 난	지금 난 기분이 너무 좋아.	굉장히 즐거우신 것 같은데요? 좋은 일이 있나요?
청년	여성	연애, 결혼, 출산	해당없음	기쁨	신이 난	나도 조카가 생겨! 너무 기뻐.	정말 좋은 소식이네요. 축하할 특별한 계획이 있을까요?
중년	여성	직장, 업무 스트레스	해당없음	불안	불안	거래처와의 다음 계약이 무산될까봐 불안해.	많이 걱정되시겠어요. 계약이 무산될 것 같다고 생각하는 이유가 있으신가요?
노년	남성	대인관계	만성질환 유	슬픔	염세적인	당뇨로 고생 중인데 아내가 이혼하자더군. 아내가 날 버리면 병은 더 악화될 텐데 세상에 나 혼자 남은 기분이야.	몸이 아픈데 아내가 이혼하자고 하니 혼자 있는 것 같은 기분을 느끼셨군요.
노년	남성	재정	만성질환 유	슬픔	마비된	노년에 재정적으로 이루어 놓은게 없어서 걱정이야.	나이들어 재정적으로 어려워서 걱정되시는군요.
노년	여성	대인관계	만성질환 유	불안	조심스러운	요즘 코로나 때문에 김 할아버지를 만나러 가지 못하고 있어.	그렇군요. 많이 슬프시겠어요.
청년	남성	대인관계	해당없음	당황	혼란스러운	하루 만에 부서 동료가 모두 다른 사람으로 바뀌어서 당황스러웠어.	부서 동료가 모두 바뀌었다니 정말 당황했겠어요.
청소년	여성	가족관계	해당없음	상처	괴로워하는	오늘 오후가 나 때문에 얼마랑 싸웠어.	정말요? 많이 속상하셨겠어요.
청년	남성	대인관계	해당없음	슬픔	슬픔	최근 물난리로 인해 직장동료가 죽었어. 다들 슬픔에 잠겨 있어.	직장동료분이 돌아가셔서 주위가 슬픔에 잠겨있군요.
노년	남성	건강, 죽음	만성질환 유	분노	구억질 나는	아내가 검사를 받고 싶다고 했는데 약국에서 약만 사주다가 결국 큰 병으로 만들고 말았어.	아내에게 큰 병이 생겨 마음이 괴로우시겠어요.
중년	여성	직장, 업무 스트레스	해당없음	분노	악의적인	박 과장님이 나에게 너무 모욕적인 말을 했어.	앞에서 모욕적인 말을 들어서 화나셨겠어요.

<감성 대화 말뭉치 데이터셋>

발표 할 때, 김종욱 교수님이 챗봇의 응답이 정확하게 나오는지 어떻게 아느냐, 평가 지표가 있는 것이냐라고 물어보셨습니다. 정확도가 82%가 나왔지만, 객관적인 외부 평가를 위해서 **KLUE**(자연어 처리 이해 평가 데이터셋)라는 데이터 셋을 이용하려고 합니다. 하지만, 현재는 데이터셋이 발표되지 않아서, 경진대회 전까지 제공되지 않는다면 다른 평가 지표를 찾아보겠습니다.

01. 개선/보완점

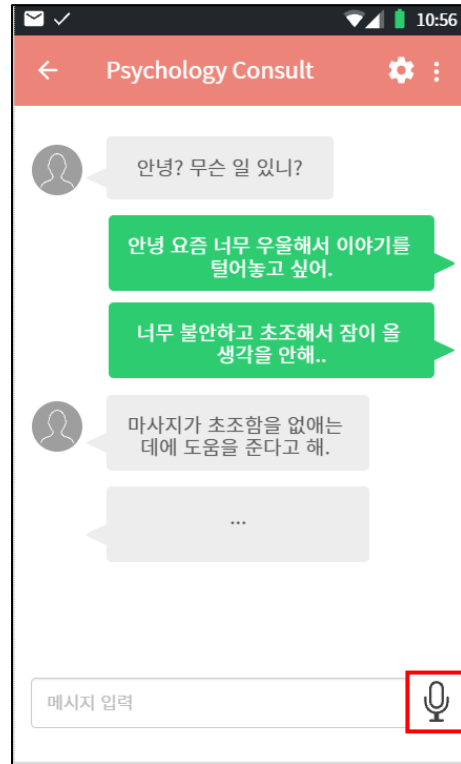
음성 합성

Multi speaker를 사용해서 여러 명의 음성을 학습하는 경우, 새로운 발화자의 음성을 학습할 때 필요한 음성 데이터가 많이 줄어듭니다. 경진 대회 전까지는 최대한 Multi speaker를 적용해볼 계획입니다.

하지만, Multi speaker 모델에 Open Data를 학습 시켜보았을 때, 아직 좋은 성능을 내는 하이퍼 파라미터를 찾지 못해서인지, Open Data의 품질 때문인지 잘 합성되지 않는 문제가 있어서 적용 가능성은 낮아 보입니다.

01. 개선/보완점

음성 인식



음성으로 챗봇과 대화하면, 대화한다는 느낌을 좀 더 강조할 수 있을 것 같아서 음성 인식 기술(Speech To Text)도 적용하려고 합니다. API는 react-native-speech-recognition이나 구글 Speech를 이용할 계획입니다.