안녕하세요. 2021년 캡스톤 디자인 컴퓨터과학과 너목들 팀 발표 시작하겠습니다.

(넘기기)

이번 발표에서는 프로젝트 소개, 구조 및 기능, 결과, 마지막으로 기대효과를 말씀드리고 발표를 마무리하겠습니다.

(넘기기)

먼저 프로젝트 소개입니다.

저희는 공감과 위로를 해주는 챗봇을 만들었습니다.

속마음을 털어 놓을 수 있는 나만의 앱, 친근한 목소리의 챗봇과 대화함으로써 마음의 위안을 얻고 대화 후 변화된 감정을 일기에 기록할 수 있는 앱 ,

나의 감정 통계를 보며 자가 진단을 할 수 있는 앱을 만들었습니다. 그런 의미에서 저희는 앱의 이름을 “들어줄게”로 이름 지었습니다.

(넘기기)

코로나 팬데믹이 길어지면서 사회적 고립감과 건강 염려 등의 이유로 성인의 40.7%가 코로나 블루를 경험했습니다.

하지만 우울증 환자에 대한 사회적 편견과 상담 시설 방문 거부감, 불이익 걱정으로 전문가에게 상담 받기에는 사회적 장벽이 높게 느껴져 52%만이 치료를 받았습니다.

따라서 사람들은 쉽게 속마음을 털어놓을 곳이 필요하고 이에 도움이 되고자 저희는 공감과 위로를 해주는 챗봇을 만들었습니다.

(넘기기)

다음으로 구조 및 기능입니다.

프론트엔드에서는 리액트 네이티브를/ 백엔드에서는 스프링 부트와 장고를 이용했습니다. 스프링 부트의 경우 앱 전반에 사용할 수 있는 API를 제공하며, 장고는 모델 학습 및 inference 등 AI 관련 API를 제공합니다.

(넘기기)

들어줄게 앱에 사용된 주요기술은 3가지로 챗봇, 어체 변환, 음성 합성입니다.

(넘기기)

먼저 음성 합성의 구조 및 기능에 대해 말씀드리겠습니다

Neural Speech Synthesis는 2 stage pipeline으로 구성되는데요, 먼저 Text로부터 Mel-Spectrogram을 예측하고, Mel-spectrogram으로부터 waveform audio를 예측하는 구조입니다. 여기서 Glow-TTS는 Flow 기반 생성 모델로, Mel-spectrogram을 예측하는 역할을 수행하며, 정렬 모델을 구축하지 않아도 보다 더 정확하게 추정하고, 딥마인드가 발표한 Tacotron2보다도 합성 속도가 빠릅니다. hifi-gan은 audio를 예측하는 역할을 수행하며, 월등히 좋은 품질의 오디오를 빠르게 합성합니다.

저희는 12시간 분량의 여성 성우의 한국어 음성 데이터인 kss 데이터 셋과 각각 2시간, 3시간 분량의 팀원 목소리를 이용하여 모델을 학습시켰습니다. 먼저 원음성을 들어본 다음 합성된 음성을 들어보겠습니다. 먼저 kss의 원음성입니다. kss의 합성음성입니다. 다음으로 제 목소리의 합성 음성입니다. 다음으로 팀원인 희진님의 원음성입니다. 합성음성입니다.

(넘기기)

다음으로, 챗봇입니다. 학습에는 AI HUB의 감성 대화 말뭉치 데이터 셋을 이용하였고, 약 12만개의 문장을 학습시켰습니다.

챗봇은 Doc2Vec 모델을 사용하였고, Doc2Vec는 Word2Vec를 문장 단위로 확장한 문서 임베딩 기법입니다.

Doc2Vec 기법은 문장 전체에 대해 단어 k개씩 슬라이딩 해가며 단어 k개가 주어졌을 때 다음 단어를 맞추는 과정을 학습합니다. 학습 결과 정확도는 82%가 나왔습니다.

사용자와 대화하는 모습에서 챗봇이 사용자의 채팅을 이해하고 상황에 맞게 응답하는 것을 볼 수 있습니다.

(넘기기)

다음으로, 저희 팀은 사용자가 챗봇과 편안하게 대화할 수 있도록 어체 변환 기술도 구현하였습니다. 모델은 Attention Encoder-Decoder 모델을 사용했고, 약 12000개의 문장을 학습하였습니다.

학습을 위해 하나의 통합 모델에 문장들 전체를 모델링하는 nmt를 이용했고, 한국어 형태소 구문 분석기를 통해 입력된 데이터를 분석한 뒤 어체를 변환시키는 구조입니다.

이미지에서 존댓말 어체를 넣었을 때 반말 어체로 변환이 되는 것을 확인하실 수 있습니다.

학습 결과 정확도는 89%가 나왔습니다.

(넘기기)

다음은 앱의 UI와 함께 주요 기능에 대해 설명 드리겠습니다.

먼저 회원 가입 과정을 통해 사용자 정보를 저장하고, 로그인 과정을 통해 사용자를 인증합니다. 합성된 음성이 다른 이용자에 의해 사용되는 것을 막기 위해 인증 과정을 넣어주었습니다.

이어지는 홈 화면은 회원가입 직후 초기화면으로, 감정 일기를 작성한 후 확인할 수 있습니다.

(넘기기)

다음은, 채팅 화면입니다.

채팅방 설정에서 대화 상대의 프로필 사진과 이름, 어체, 음성을 설정할 수 있습니다.

그렇게 생성된 채팅방 목록을 확인한 후 챗봇과 대화가 가능합니다.

대화 진행 시 말풍선을 버튼을 누르면 합성된 음성이 플레이 됩니다.

(넘기기)

다음은, 음성 데이터 생성 화면입니다.

학습에 필요한 대화 상대의 음성 파일을 저장소에서 가져오거나

직접 녹음해서 업로드 할 수 있습니다.

업로드한 음성 데이터를 학습해 대화 상대의 음성을 합성할 수 있습니다.

(넘기기)

다음으로, 감정 일기입니다.

일기를 통해 생각을 정리하면 감정적으로 해소되는 듯한 느낌을 받을 수 있습니다. 이러한 기분 개선 효과와 챗봇 대화 전후의 기분 변화를 살펴보기 위해 이모지를 사용하여 현재의 감정 상태를 기록하고, 감정 일기를 작성할 수 있습니다.

작성된 감정일기는 이모지의 이름과 일기 내용으로 검색이 가능하며 일기를 선택하면 해당 일기의 상세한 내용을 확인할 수 있습니다.

(넘기기)

감정 통계는 감정 일기에 작성된 이모지를 차트화 한 화면으로 감정의 월별/일별 통계와 시간에 따른 감정의 변화를 한눈에 확인할 수 있습니다.

감정 일기 작성 후 홈화면에서는 사용자가 작성한 감정 일기를 기반으로 상단에는 가장 최근 일기에 등록된 감정 이모지가 표시되고, 하단에는 감정별 분포를 차트로 보여줍니다.

(넘기기)

다음으로 프로젝트 결과인 시연 영상을 보도록 하겠습니다.

(넘기기)

마지막으로 기대효과입니다.

첫번째로, 언제 어디서나 위로 받고 공감 받을 수 있어 시간적, 공간적 제약이 극복될 수 있습니다.

두번째로, 맞춤 음성 설정이 가능해 지인이 공감해주는 듯한 편안한 대화가 가능합니다.

세번째로, 상담 비용으로 인한 부담이 해소되고 사생활 공개의 걱정이 극복되어 이용자 부담이 완화될 수 있습니다.

(넘기기)

이상 너목들 팀 발표를 마치도록 하겠습니다. 감사합니다.