교수님 요구사항

Point

1. 노인에게 질문 던지기

2. 상황파악(IOT, 센서)

3. 챗봇이 스스로(active)

4. 메뉴기반 보다는 음성기반으로

5. photo alive 열심히 사용해보고 챗봇으로 어떻게 활용할지 생각해보기(노인 말벗)

6. 구글 speech로 TTS,STT서비스

7. AI로 상황인식

-Watson IBM api등을 사용 / 어떤 기능이 있는지

-필요한 물품,S/W,H/W(상품 종류, 범위,가격) 그 외에 문의해보고 기록, 알아보기

8. 시간 내에 할 수 있는게 중요

9. AI, IOT, ChatBot 기존에 있는 것 찾아보고 +a로 api, 서버연동 공부

+알파

1. 대화와 공감이 필요(1문 1답의 채팅 보다는 multi-turn을 기본으로 하는 대화형)

2. 메뉴 base 말고 active하게

3. 너무 많은 기능을 넣기 보다 제한된 범위, 좁은 주제에 대한 가치 제공, 잘 해내야함

4. 불필요한 정보를 제공하면 안됨.(끄라했을 때 적절히 끄는게 중요(명령에 대한 이해))

5. 자연어 처리가 중요!(다양한 data들이 자연어 형태로 존재)

6. 말꼬리를 이어 나가는것보다 공감하고 맞장구가 낫지 않을까?

7. 누가 발화할 차례인지 판단이 필요. 일단은 상대방의 말의 템포가 끝나면 발화하는 것으로 설정

8. 틀린 정보를 말하느니 안말하는게 나을 듯 -> 단순화

9. Response with context-aware need to generate sentence not only from the ask, but with context-aware data, knowledgebase, decision making process

Idea

1. 챗봇을 이용한 건강 check

-취득한 정보를 바탕으로 가족이나 주치의 등에게 알림 or rule base로 ‘아픔’을 의미하는 말뭉치 dictionary를 만들고 관련된 말이 들어왔을 때 유사도를 계산하여 가장 유사도가 높은 말에 대응하는 대답해주기.

-아픈 곳들 별로 병원을 나누고 위치 기반 api를 사용하여 근처 병원을 검색하여 줌

1)’아픔’을 의미하는 말뭉치

-국립국어원에서 말뭉치 뽑아내고, 아프리카, 아프간, 아파트 등을 제거하여 추출

-유사도 구할 때 tf-idf 활용

2)아픈 부위별 병원 나누기

-rule base

3)위치기반 API(구글) 찾고 활용 방법

4)근처 병원 검색해주는 방법

2. facebook에 가족들이나 미리 등록한 사람이 게시글을 올렸을 때 챗봇이 노인에게 말을 걸어 알려주고 음성서비스로 댓글이나 메시지를 남길 수 있도록

-알림 : 보시겠습니까? -> 응 보여줘 vs 아니 괜찮아 -> ‘응’일 경우 : 메시지를 보내시겠습니까?

-> photo alive 활용

3. 노인의 감정분석 (with NLP) -> 감정을 공감하는 말을 해주면 놀랄 듯

방법

1. rule base

-keyword가 있을 경우 응답

2. Q/A (IBM- WatSon)

-방대한 text에서 답을 찾아서 제공하는 지능형 검색 시스템

-유사도 기반 검색 방법

1)대화쌍 DB구축

->최대한 많이, 잘(노인이라는 대상에 맞도록!)

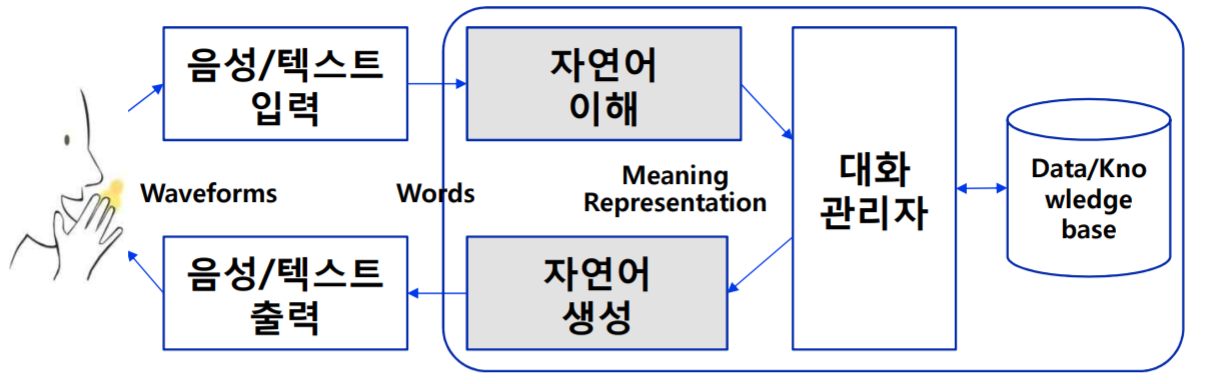
->기본적 인사말, 확신이 없을 때 적당히 사용할 수 있는 상황별 응답 대화, 영역 서비스를 위한 info-chatting용 대화쌍

2)사용자의 발화와 가장 유사한 대화쌍 찾기

-한계 : 유사한 발화가 없다면?

3. Info-chatting

-관련 정보를 관련 DB나 관련 시스템으로부터 찾아서 응답하는 채팅



-자연어 처리의 정확도, 응답 DB의 콘텐츠, 대화 generator의 성능?

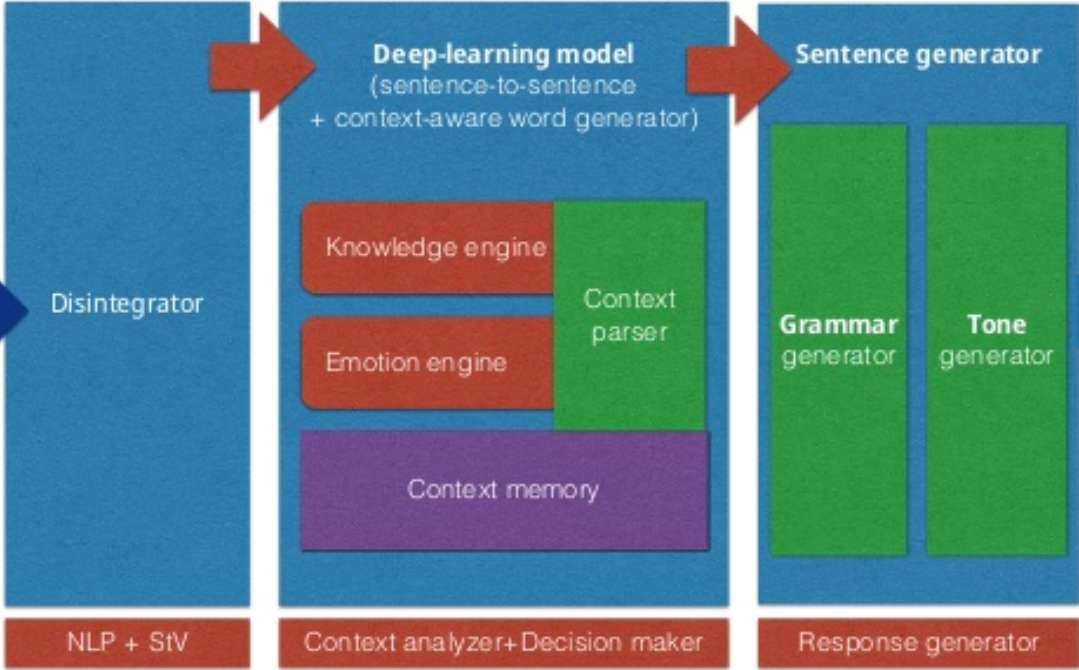
4. 딥러닝

-hierarchical seq2seq model for generating dialogues

-많은 labeling된 training data필요

-unlabeled data 자동 clustering 능력을 이용하여 새로운 class발견 가능성도 있음

-강화학습을 통한 decision making



\*Disintegrator

-모델 그냥 쓰면됨(KoNlP나 twitter나 komoran..?)

\*Deep-learning model – Context Parser

1) knowledge engine

2) Emotion engine

-senti-word-net to extract emotion

-감정을 6개정도의 범주로 나누기

-텍스트를 벡터로 변환 후(한 문장을 구성하는 단어들을 word2vec으로 돌린 후 나온 값을 CNN에) CNN에 넣은 후 max pooling해서 나온 feature들을 LSTM에 넣기(+attention도..?)

\* Sentence generator

1) Grammar generator model

-input : 명사, 동사, 형용사 등의 sequence를 input으로 넣음(postpositions, 접속사 등 없는채로)

-output : sequence of normal / monotonic sentence

-> 일반 문장을 많은 sequence로 쪼갬

->postpositions, conjunction 제거

->명사, 동사, 형용사를 normalize

->seq2seq 모델에 (각 encoder, decoder에) 3 layer 정도 학습

->hidden feature size of GRU cell 수 : 25, Embedding dimension for each word 개수 : 25

2) Tone generator model – 사용자의 선호에 따라 다양한 톤으로(더 유용함)

-더 sentence를 사람처럼(형태소 분석)

\*Training Bot Model

1) Input

-Disintegrated sentence sequence without potpositions/conjunction

-Emotion flag

-Context flag

2) Output

-Answer sequence with nouns, pronouns, verbs, adjectives

3)Learning

-Supervised learning(for simple communication model)

-Reinforcement learning(for emotion, context flag)

5. 챗봇이 먼저 대화를 걸도록

How?

TV 켜져있을 때 일정 시간 대화가 없으면 -> 질문

언제 어디서 무엇을 어떻게 왜 누구와

가족들의 음성메모를 이용한 QA

생성..? 단어들을

조사를 다 해보고 해도 안나오면 뭐..

API

우선 첫 출근 후 1시간반 정도 동안 앞으로 어떤일을 할지 주제를 설명받았습니다. 큰 주제는 active한 노인 대상 챗봇을 구현하는 것이며 그에 대한 요구사항은 다음과 같습니다. 1. 노인에게 질문 던지기, 2. 상황 파악, 3. 챗봇이 스스로 active하게 말하도록, 4. 음성기반으로 5. 미리 구현되어 있는 'photo alive'라는 기술 사용하여 어떤식으로 챗봇에 활용할지 생각해보기 6.사용할 API 찾아보기 7.AI로 상황인식, 8. 필요한 것들 정리하기 입니다.

따라서 1주일 동안 어떤식으로 구현할지, 어떤 기능을 추가할지 등을 조사하고 기획하여 다음주 월요일 아침에 발표하기로 하였습니다. 따라서 우선 처음 다루어보는 챗봇이기에 어떤 API가 있는지 찾아보았고 지금까지 있던 passive한 챗봇과는 다르게 어떻게 active한 챗봇을 개발할지도 논문을 찾으며 공부하였습니다.

오늘은 어떤 기능을 구현할지 정했습니다. 제 아이디어는 크게 3가지이고 다음과 같습니다. 1.챗봇을 이용한 건강check이며 취득한 정보를 바탕으로 가족이나 주치의 등에게 알림이 가도록 하는 것입니다. rule base방식으로 말뭉치를 만들어놓고 들어오는 말과 비교하여 내뱉는 방식입니다. 2. facebook에 가족들이나 미리 등록한 사람이 게시글을 올렸을 때 챗봇이 노인에게 말을 걸어 알려주고, 음성서비스로 메시지를 남길 수 있도록 하는 기능입니다. 이것또한 rule base로 말뭉치를 만드는 방식으로 생각했습니다. 3. 노인의 감정분석입니다. 자연어처리를 공부했기에 말을 text로 바꾸어 감정을 분석하고 노인의 감성을 인식하는 것입니다. 이렇게 3가지를 생각한 후 관련 논문을 찾아보았고 딥러닝을 위해 많은 데이터가 필요하므로 어디서 데이터를 구할지 알아보았습니다. 내일은 실질적인 구현과 API사용법에 대하여 알아볼 생각입니다.