

와이즈넛 인공지능 NLP 기술 소개

챗봇 기술

성장기술연구소 장정훈 연구소장



주인공 토니스타크를 도와주는 인공지능 집사 Jarvis

[스타크]

"자비스, 심장 체크해봐. 내 뇌가 이상한거야?"

[자비스]

"심장이랑 뇌에는 전혀 문제가 없어요."

[스타크]

"그럼, 독극물에 중독됐어?"

[자비스]

"제가 보기엔 심각한 불안증세예요."

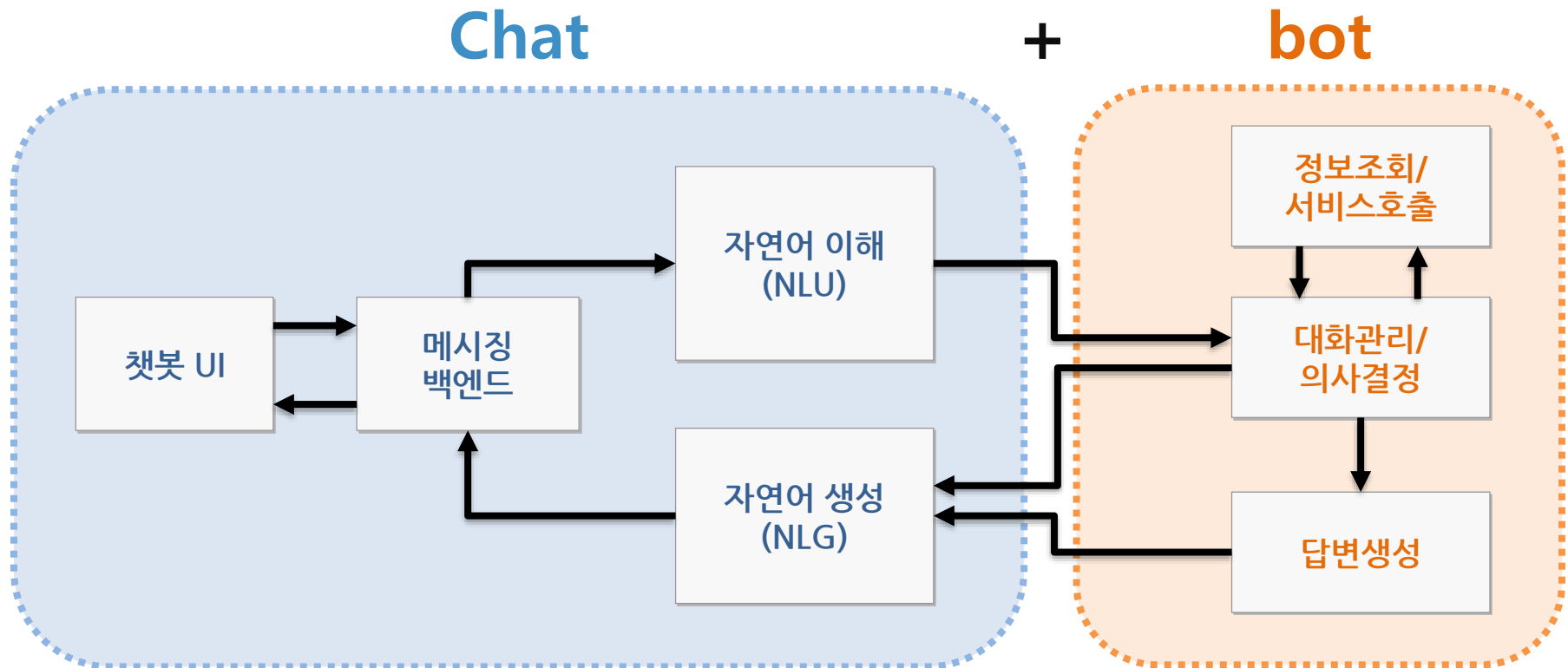
1

챗봇

1.1 챗봇이란?



챗봇(chatbot)은 음성이나 문자를 통한 **인간과의 대화**를 통해서 **특정한 작업**을 수행하도록 제작된 컴퓨터 프로그램이다¹⁾



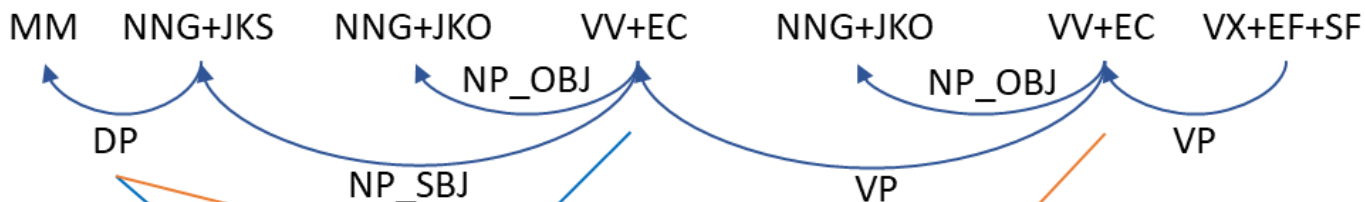
¹⁾ ["What is a chatbot?"](#). techtarget.com. Retrieved 30 January 2017.

Gartner Hype Cycle for Artificial Intelligence, 2021



1.2 자연어 처리(Natural Language Processing)

한 남자가 스마트폰을 보며 횡단보도를 건너고 있다.



형태소 분석

의존구문 분석

용언: 보다
행위주: 한 남자
피동작주: 스마트폰

용언: 건너다
행위주: 한 남자
피동작주: 횡단보도

의미역 인식

SUB: 한 남자
OBJ1: 스마트폰, OBJ2: 횡단보도
보다 (SUB, OBJ1)
건너다 (SUB, OBJ2)

위험하다 (SUB)

의미 분석 & 추론

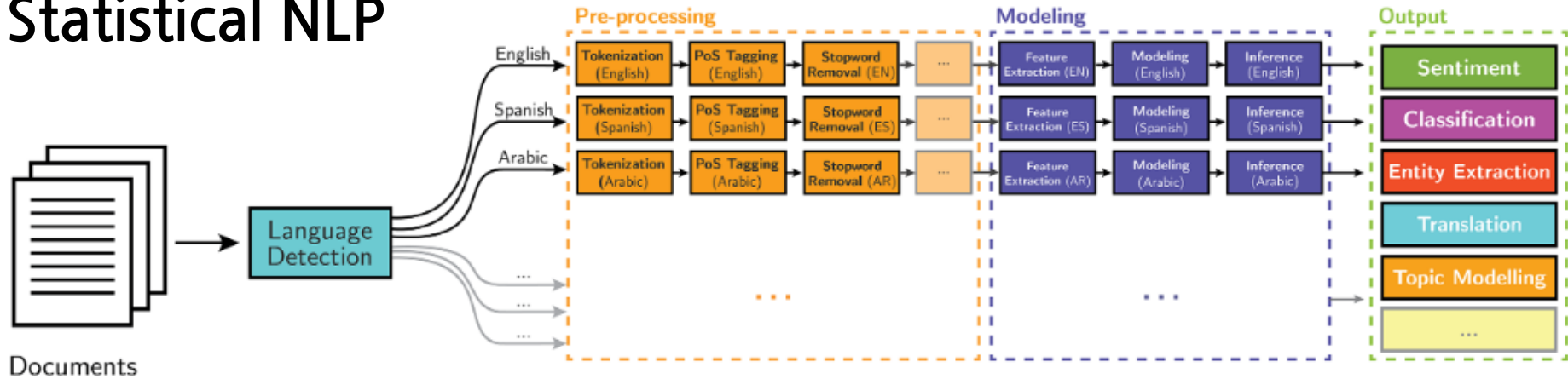
화자는 언어처리의 과정과 예제를 보여주기 위해
이 문장을 예로 들고 있다.

화용 분석

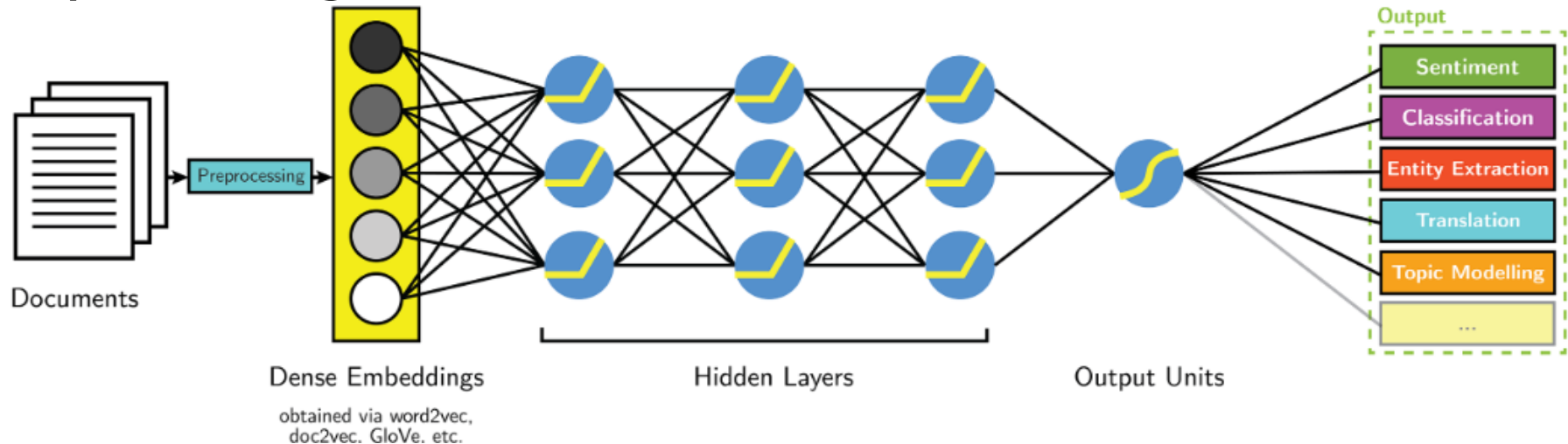


1.2 딥러닝 기반 자연어 처리(Natural Language Processing)

- Statistical NLP



- Deep-learning NLP

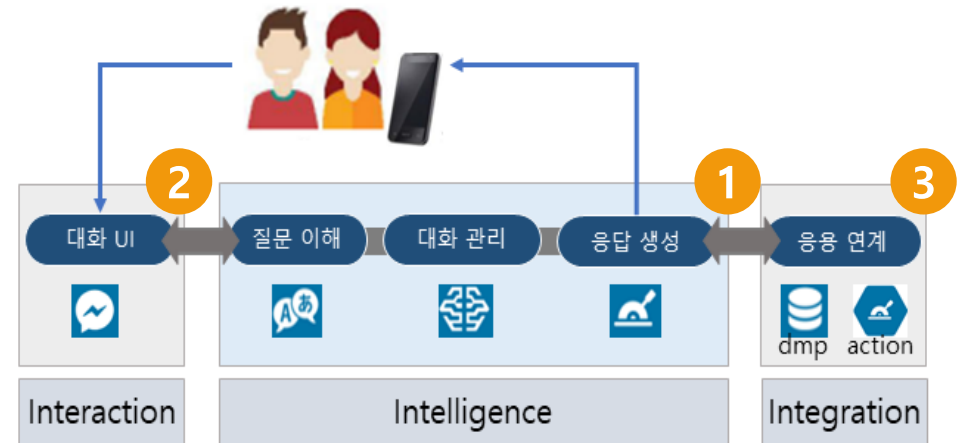
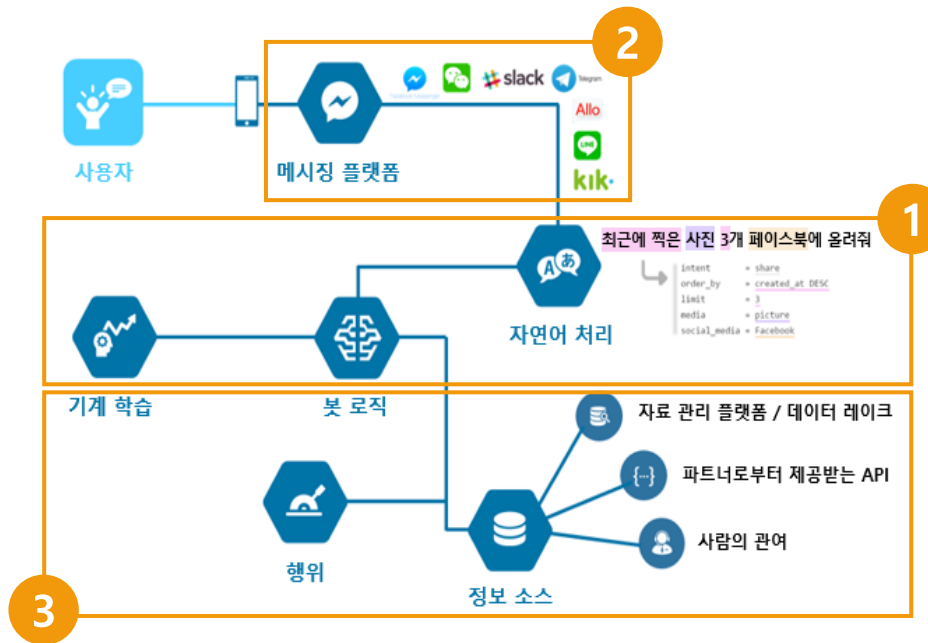


1.3 AI 챗봇 구현에 필요한 기술과 기능 구성



사람들마다 챗봇의 기대수준 다양

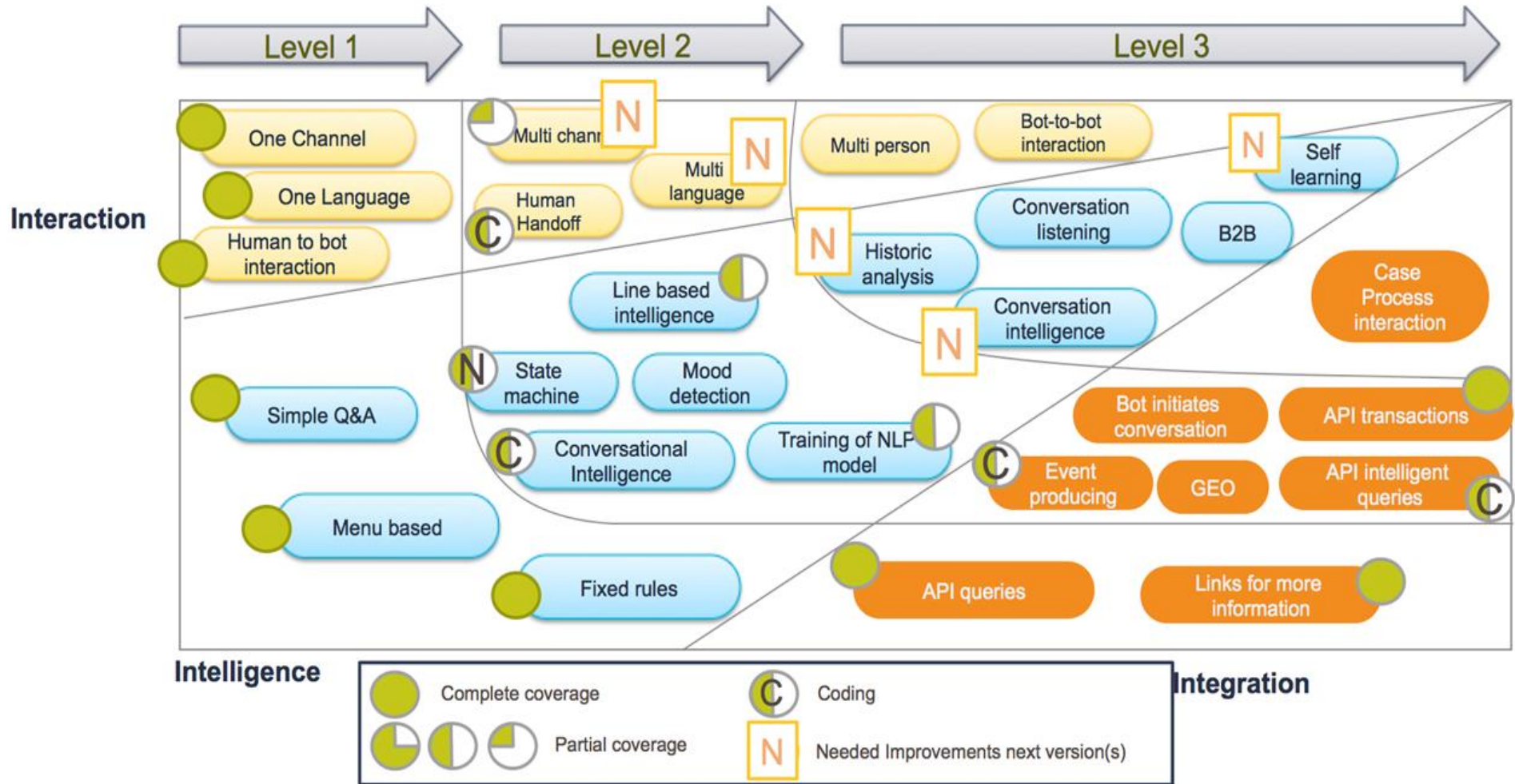
- 특정 주제에 대하여 미리 작성해 놓은 답변을 하는 챗봇
- 모든 주제에 모두 답이 가능한 챗봇까지



* 출처 : Chatbotsmagazine, "The Ultimate Guide To Designing A Chatbot Tech Stack", Sebastien Fourault, 2017

1.4 AI 챗봇 기술 맵과 성숙도

챗봇의 각 구성 별 단계적인 기술 수준



* 출처 : <https://www.Capgemini.com/2017/09/oracle-goes-chatbot/>

1.5 Intelligence>질문이해

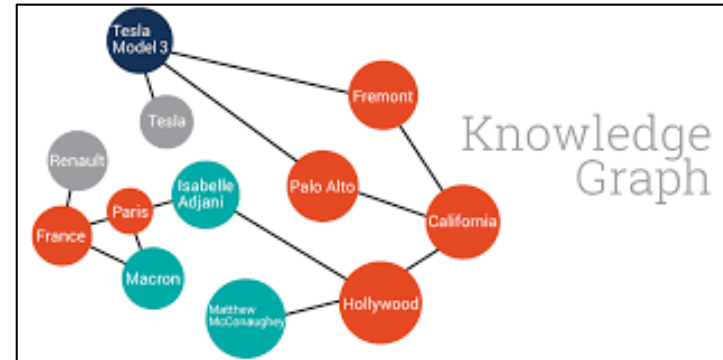
1. Keyword, Pattern Matching

- Q&A, 키워드 검색 기반



2. QA Knowledge Graph

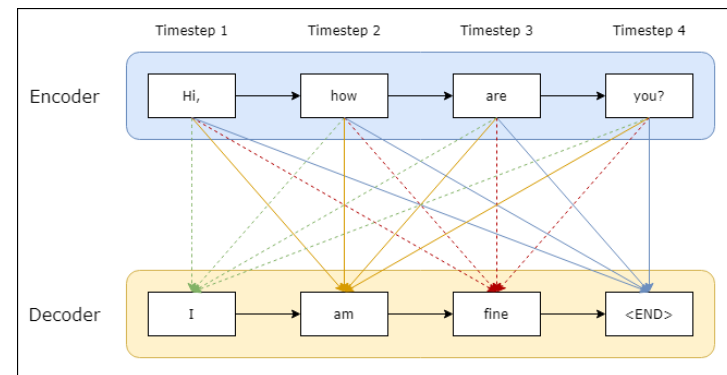
- 온톨로지 기반
- 질문 키워드 매칭, 추론



3. Intent, entity, Classification

contentSkip to site indexPoliticsSubscribeLog InSubscribeLog InToday's PaperAdvertisementSupported ORG by F.B.I. Agent Peter Strzok PERSON , Who Criticized Trump PERSON in Texts, Is FiredImagePeter Strzok, a top F.B.I. GPE counterintelligence agent who was taken off the special counsel investigation after his disparaging texts about President Trump PERSON were uncovered, was fired. Credit T.J. Kirkpatrick PERSON for The New York TimesBy Adam Goldman ORG and Michael S. SchmidtAug PERSON 13 CARDINAL , 2018WASHINGTON CARDINAL — Peter Strzok PERSON , the F.B.I. GPE senior counterintelligence agent who disparaged President Trump PERSON in inflammatory text messages and helped oversee the Hillary Clinton PERSON email and Russia GPE investigations, has been fired for violating bureau policies. Mr. Strzok PERSON 's lawyer said Monday DATE Mr. Trump and his allies seized on the texts — exchanged during the 2016 DATE campaign with a former F.B.I. GPE lawyer, Lisa Page — in PERSON assailing the Russia GPE investigation as an illegitimate "witch hunt." Mr. Strzok PERSON , who rose over 20 years DATE at the F.B.I. GPE to become one of its most experienced counterintelligence agents, was a key figure in the early months DATE of the inquiry. Along with writing the texts, Mr. Strzok PERSON was accused of sending a highly sensitive search warrant to his personal email account. The F.B.I. GPE had been under immense political pressure by Mr. Trump PERSON to dismiss Mr. Strzok PERSON , who was removed last summer DATE from the staff of the special counsel, Robert S. Mueller III PERSON . The president has repeatedly denounced Mr. Strzok PERSON in posts on

4. Sequence to sequence



1.6 ELMO, BERT

ELMO



(**E**MBEDDING for **L**ANGUAGE **M**ODEL)

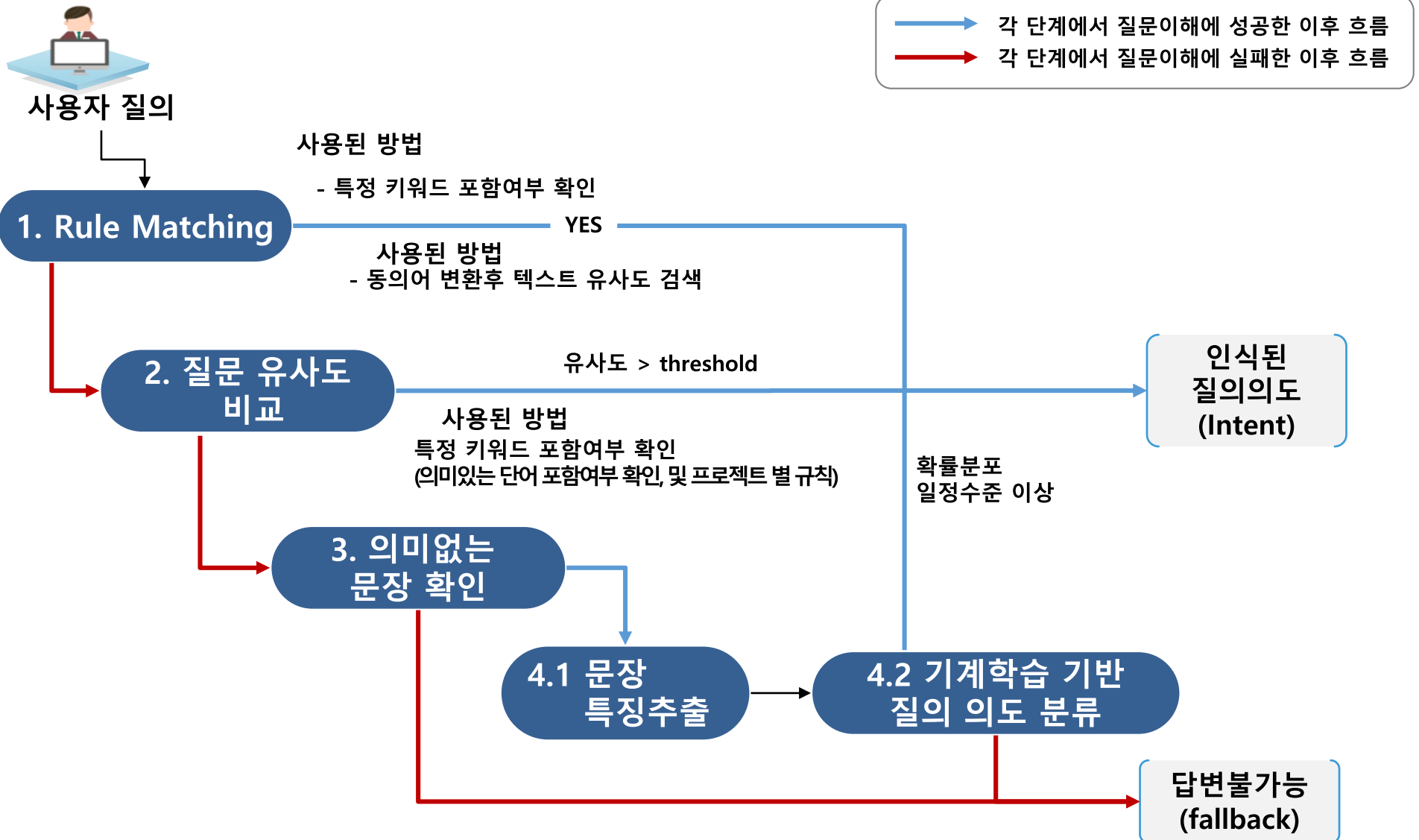


BERT



(**B**IDIRECTIONAL **E**NCODER **R**EPRESENTATIONS
from **T**RANSFORMERS)

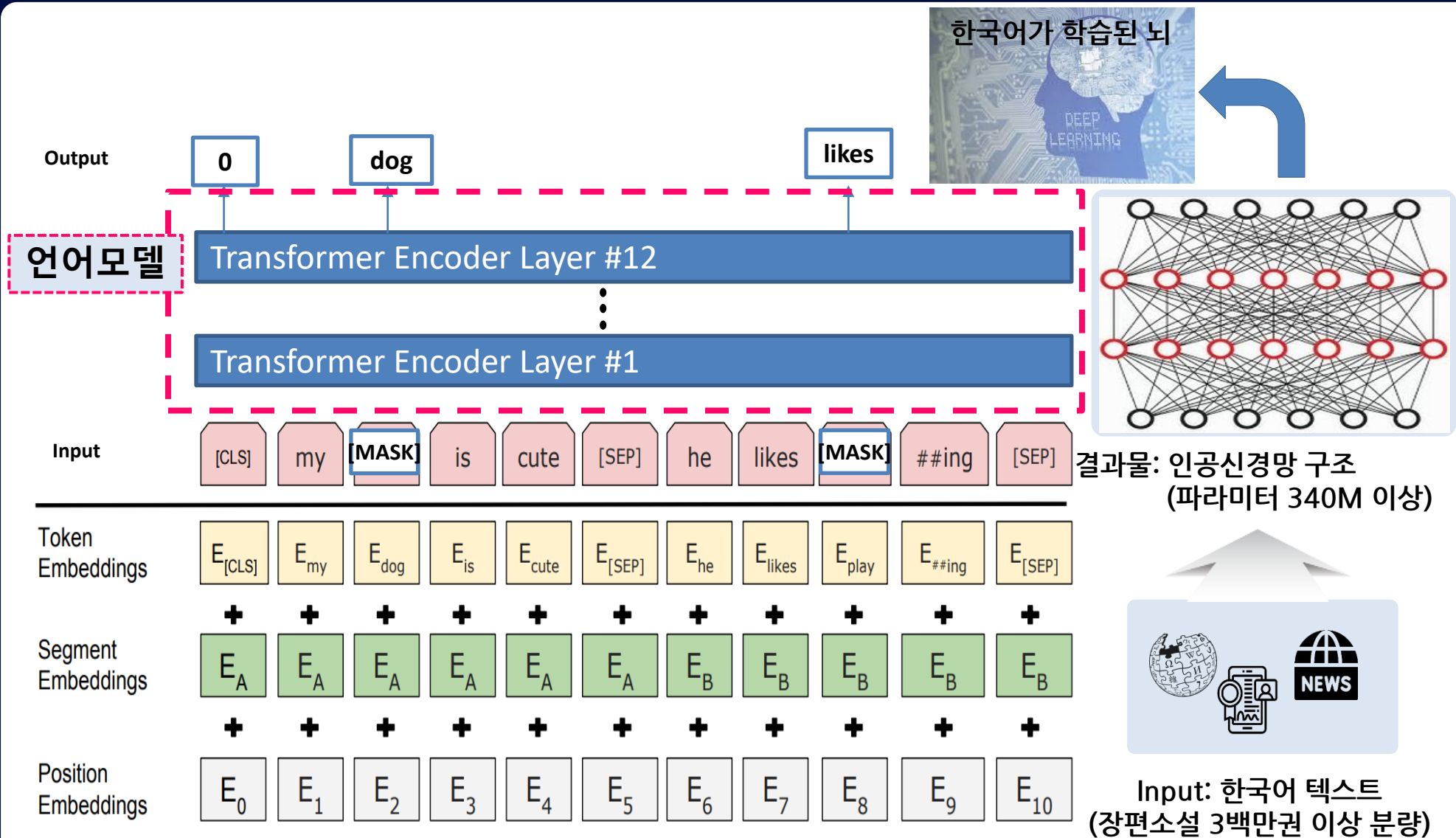
1.7 와이즈넷 챗봇 – 질문 이해 흐름도



2

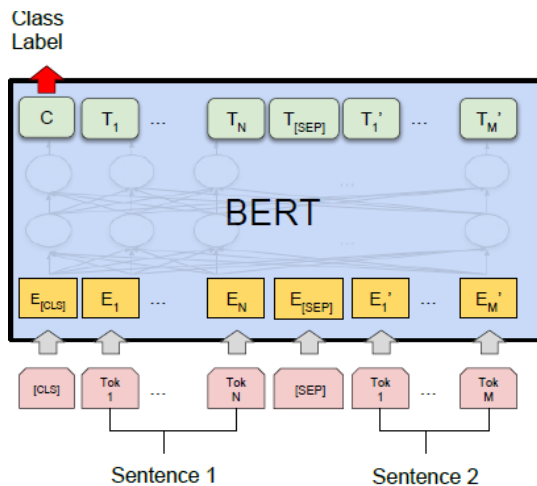
언어모델

2.1 사전 학습 언어 모델(Pre-trained Language Model)



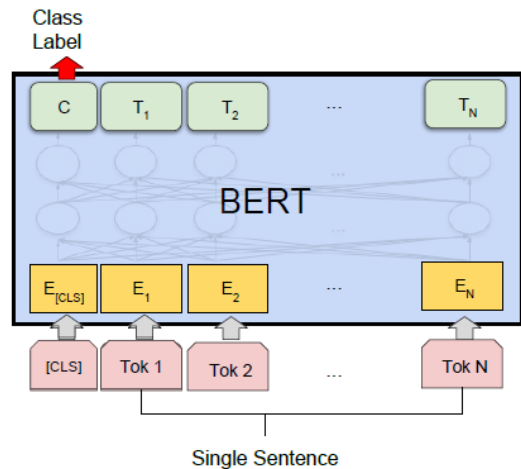
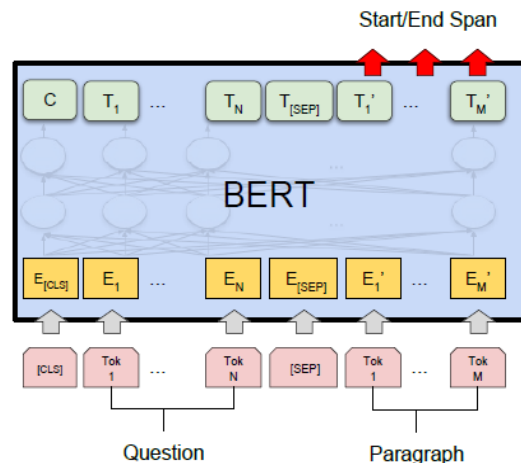
2.2 언어 모델을 이용한 다운스트림 태스크

- 사전 학습된 모델을 지도학습을 통해 특정 태스크에 맞게 미세 조정(Fine-tuning)
- 동일한 사전 학습 모델을 다운스트림 태스크를 통해 여러 문제에 적용 가능



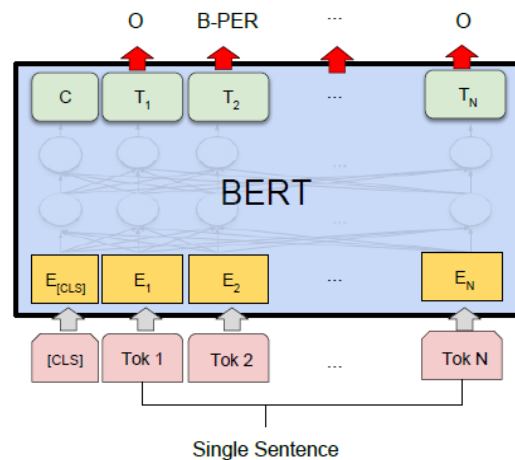
문장 쌍 분류
(의미적 추론)

구간 선택
(기계독해)



단일 문장 분류
(감정분석)

시퀀스 태깅
(개체명 인식)



3

3. 언어 모델과 '솔루션 챗봇' 기술 방향

3.1 솔루션 챗봇 기술 방향

1. 상담원 협업 도구

- 주제에서 벗어난 응답을 하는 경우 상담원에게 알림
- 고객이 챗봇 대화 상담에 불편을 느끼거나 반복된 질문을 하는 상황 알림
- 대화 내용 및 매뉴얼 요약

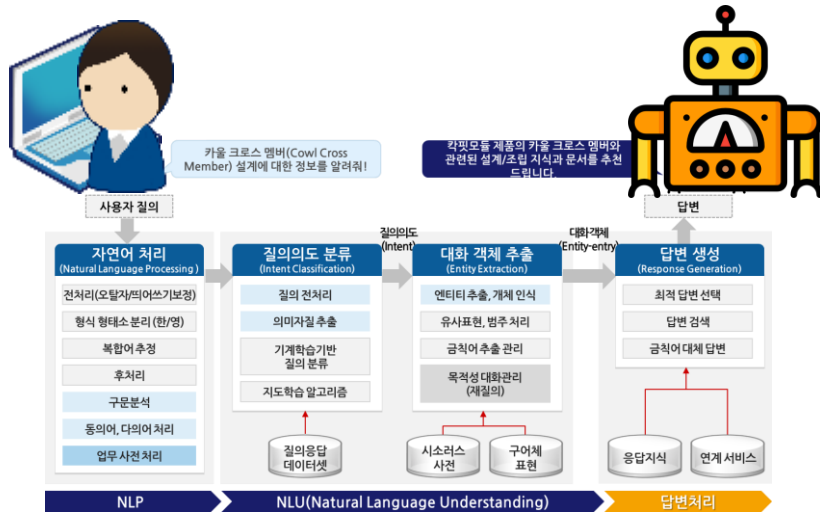
2. 멀티턴 대화 이해 및 텔레마케터 대화 도우미

- 효율적인 텔레마케팅을 위한 성공 케이스 대화 추천
- 직전의 질문만 이해하는 것이 아닌 대화 맥락을 이해하고 최적 답변 추천

3. 기계 독해

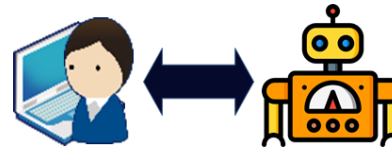
- 내부 규정이나 매뉴얼 내용을 이해하고 질문에 답변

3.2 상담원 협업 도구



대화 내용 모니터링

- 정상적인 인텐트 인식 되었는가?
- 답변이 잘 되고 있는가?
- 대화 내용 요약
- 주제에서 벗어나고 있는가?



이상한 상황 알림

- 고객이 똑같은 질문을 계속하는가?



고객 불만 포착

- 고객이 화를 내는지 파악
- 고객이 상담원을 요청하는가?



3.2 상담원 협업 도구 – 이상한 상황 알림

최종
목표

이전 대화 내용을 분석하여 주제에서 벗어나는지 감지

✓ 특허 출원 및 관련 논문 게재

챗봇 대화 시스템에서 대화 주제를 매칭하고
시스템에 알람을 하는 방법 (특허출원, 2021/12/09)

2021-12-09	
【서지사항】	
【서류명】	특허출원서
【참조번호】	2021P1855
【출원구분】	특허출원
【출원인】	
【명칭】	주식회사 와이즈넷
【특허고객번호】	1-2000-031434-3
【대리인】	
【명칭】	특허법인임앤정
【대리인번호】	9-2017-100221-7
【지정된변리사】	임승섭, 정우성, 윤락근
【포괄위임등록번호】	2017-059269-2
【발명의 국문명칭】	챗봇 대화 시스템에서 대화 주제를 매칭하고 시스템에 알람 하는 방법
【발명의 영문명칭】	THE METHOD OF MATCHING DIALOG TOPICS IN THE CHATBOT CHATTING SYSTEM AND OF ALERTING THE SYSTEM

【발매지】

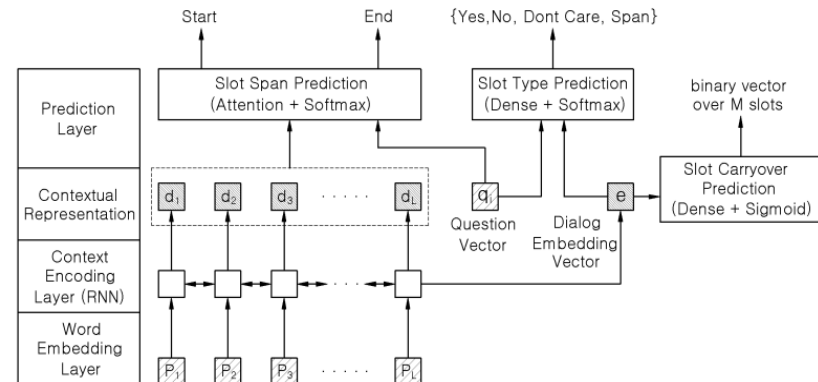


formation

will be held at AAAI-22, Dates: TBD

✓ 내용

1. 각 대화의 주제 label을 정하고 현재 대화의 주제 벡터와
코사인 유사도 계산을 통해 비교
2. 챗봇 및 사용자 대화 상태 추적
3. 두 기술을 통해 챗봇이 엉뚱한 대답을 하는지 파악하여 알람



3.2 상담원 협업 도구 - 대화 내용 요약

- T5 모델을 활용하여 대화 내용 요약
- 모델의 입력(화자 분리)을 자연스러운 문장으로 변경하여 입력



화자1 : 어디예요?
화자2 : 오산이응.
화자2 : 집가서 일 해야될 것 같아쇼.
화자1 : 병원가려구?.
화자2 : 바로 집으로 가게.
화자1 : 헐...
화자1 : 안과는못가구?.
화자2 : 응.
화자2 : 가서 급하게 뭐 해야댕



정답(by human)

안과는 못 가고 **집에 가서** 급하게
일을 해야 할 것 같다.

예측(output of model)

일을 해야 할 것 같아서 바로
집으로 가려고 한다.

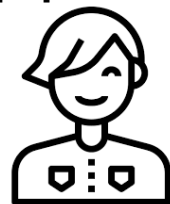
모델	ROUGE1, ROUGE2, ROUGEL
kogpt2	26.28, 9.70, 25.06
t5-base(ours)	49.697, 25.401, 37.549

3.2 상담원 협업 도구 - 대화 내용 요약

정녕 이것이 서애 영감께서 보내신 것이란 말이나?
조정이 저들과 교린을 하지 않겠다니. 강하게 맞서려
할수록 저들은 더욱강성해질 것임을 왜 모른단
말인가. 아무리 천리가 떨어져 있다 하나모두 전하의
백성들이거늘,그 목숨에 어찌 이리도무심할 수 있단
말인가? **화자2**]다 들었시요. 이제는 알았음메. 만호
나리가 말한 방법으로는 안 됨메. 이제는 우리 식대로
하겠어요. **죽을 때 죽더라도 한 번 싸워나 보고
죽겠습니다**. 화자1]망동하지 마라. 화자2]**내
못하겠습니다**. 그 이러다 한 번 싸워보지도 못하고
애미나이를 잃고 말 것 아니오. 화자3]그렇습니다,
막지 마십시오. **화자1**]명령이다, **망동하지 말고 기다려.
내 북병영으로 갈 것이니라**. 화자2]나리, 이건
싸움이고 전쟁이에요. 말로 하는 게 아니다. 이 말임메!
화자1]기다리라지 않느냐? 함부로 망동하는 자는 내
군율에 의거참하고야 말 것이야. 알겠는가?

정답 (by human)

조정의 교린 의사가 없음을 알자,
화자2,3은 싸우려고 하고, 화자
1은 이를 말린다



예측 (output of model)

조정이 저들과 교린하지 않겠다고
하자 화자2는 한 번 싸워보고
죽자고 하지만 화자1은
북병영으로 가겠다고 한다



3.3 멀티턴 대화 이해(1/3)

● 딥러닝 기반 멀티턴 대화 향상 기법

Utterance Manipulation Strategy

다이얼로그 컨텍스트로부터 다음 응답을 선택하기 위해
"발화 조작 전략"의 개념을 이용하여 성능을 개선

Utterance Insertion

Utterance Deletion

Utterance Search

Utterance Insertion / Deletion

- 연속된 원래의 대화 컨텍스트에서 연속 발화를 추출하고, 발화를 삽입, 삭제해야 하는 위치를 찾으려 학습

Utterance Search

- 무작위로 섞인 발화에서 마지막 발화의 직전 발화를 찾는 것을 목표로 학습

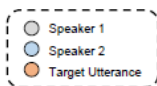
[Dialog Context]

Good morning! What can I do for you?
I'm thinking of traveling to California in May.
Could you recommend some tourist programs for that?
With pleasure. We arrange two kinds of tourist programs for California, a seven-day tour by bus and a five-day flying journey.
How much does a seven-day tour by bus cost?
Two thousand dollars.
Does that include hotels and meals?
Oh, yes, and admission tickets for places of interest as well.

[Response]

That sounds reasonable.

(a) Response Selection



[Utterance Insertion]

(a) I'm thinking of traveling to California in May.
Could you recommend some tourist programs for that?
(b) With pleasure. We arrange two kinds of tourist programs for California, a seven-day tour by bus and a five-day flying journey.
(c) [Target Utterance]
(d) Two thousand dollars.
(e) Does that include hotels and meals?
(f) How much does a seven-day tour by bus cost?

[Utterance Deletion]

(a) I'm thinking of traveling to California in May.
Could you recommend some tourist programs for that?
(b) With pleasure. We arrange two kinds of tourist programs for California, a seven-day tour by bus and a five-day flying journey.
(c) I'd like to taste some local dishes. What would you recommend?
(d) Two thousand dollars.
(e) Does that include hotels and meals?

Random Dialog

[Utterance Search]

(b) With pleasure. We arrange two kinds of tourist programs for California, a seven-day tour by bus and a five-day flying journey.
(c) How much does a seven-day tour by bus cost?
(d) Does that include hotels and meals?
(e) Two thousand dollars.
(f) I'm thinking of traveling to California in May.
Could you recommend some tourist programs for that?
(g) Oh, yes, and admission tickets for places of interest as well.

Previous Utterance

(b) Utterance Manipulation Strategies

3.3 멀티턴 대화 이해(2/3)

● AAI 2021 와이즈넷 제1 저자 등재

세계적인 권위의 국제인공지능학회(AAI) 2021에서 와이즈넷의 '딥러닝 기반 멀티턴 대화 향상 기법' 논문이 게재



전세계 컴퓨터사이언스 분야 최고 권위 학회
국제인공지능학회
"AAAI 2021 논문 채택"



전세계 기고 논문



한국 기고 논문



한국 채택 논문

딥러닝 기반 멀티턴 대화 향상 기법

PRELIMINARY VERSION: DO NOT CITE
The AAAI Digital Library will contain the published version some time after the conference

Do Response Selection Models Really Know What's Next?
Utterance Manipulation Strategies for Multi-turn Response Selection

Taesun Whang^{1*} Dongyub Lee^{2*} Dongsuk Oh³ Chanhee Lee³
Kijong Han⁴ Dong-hun Lee⁴ Saehyeok Lee^{1,3†}
¹Wisnet Inc.
²Kakao Corp.
³Korea University
⁴Kakao Enterprise Corp.

Abstract

In this paper, we study the task of selecting the optimal response given a user and system utterance history in retrieval-based multi-turn dialog systems. Recently, pre-trained language models (e.g., BERT, RoBERTa, and ELECTRA) showed significant improvements in various natural language processing tasks. This and similar response selection tasks can also be solved using such language models by formulating the tasks as dialog-response binary classification tasks. Although existing works using this approach successfully obtained state-of-the-art results, we observe that language models trained in this manner tend to make predictions based on the relatedness of history and candidates, ignoring the sequential nature of multi-turn dialog systems. This suggests that the response selection task alone is insufficient for learning temporal dependencies between utterances. To this end, we propose utterance manipulation strategies (UMS) to address this problem. Specifically, UMS consist of several strategies (i.e., insertion, deletion, and search), which aid the response selection model towards maintaining dialog coherence. Further, UMS are self-supervised methods that do not require additional annotation and thus can be easily incorporated into existing approaches. Extensive evaluation across multiple languages and models shows that UMS are highly effective in teaching dialog consistency, which leads to models pushing the state-of-the-art with significant margins on multiple public benchmark datasets.

Introduction

In recent years, building intelligent conversational agents has gained considerable attention in the field of natural language processing (NLP). Among widely used dialog systems, retrieval-based dialog systems (Lowe et al. 2015; Wu et al. 2017; Zhang et al. 2018) are implemented in a variety of industries because they provide accurate, informative, and promising responses. In this study, we focus on multi-turn response selection in retrieval-based dialog systems. This is a task of predicting the most likely response under a given dialog history from a set of candidates.

*These authors equally contributed to this work.

†Corresponding author.

Copyright © 2021, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.

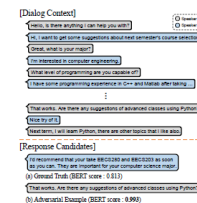


Figure 1: Example of multi-turn response selection. BERT-based model tends to calculate the matching score of a dialog-response pair depending on the semantic relatedness of the dialog and the response (a) < (b). More details are in Discussion section.

Existing works (Wu et al. 2017; Zhou et al. 2018; Tao et al. 2019a; Yuan et al. 2019) have studied utterance-response matching based on attention mechanisms including self-attention (Vaswani et al. 2017). Most recently, as pre-trained language models (e.g., BERT (Devlin et al. 2019), RoBERTa (Liu et al. 2019), and ELECTRA (Clark et al. 2020)) have achieved substantial performance improvements in diverse NLP tasks, multi-turn response selection also has been resolved using such language models (Whang et al. 2020; Lu et al. 2020; Gu et al. 2020; Humeau et al. 2020).

However, we tackle three crucial problems in applying language models to response selection. 1) Domain adaptation based on an additional training on a target corpus is extremely time-consuming and computationally costly. 2) Formulating response selection as a dialog-response binary classification task is insufficient to represent intra- and inter-utterance interactions as the dialog context is formed by concatenating all

3.3 멀티턴 대화 이해(3/3)

성능 참고 자료

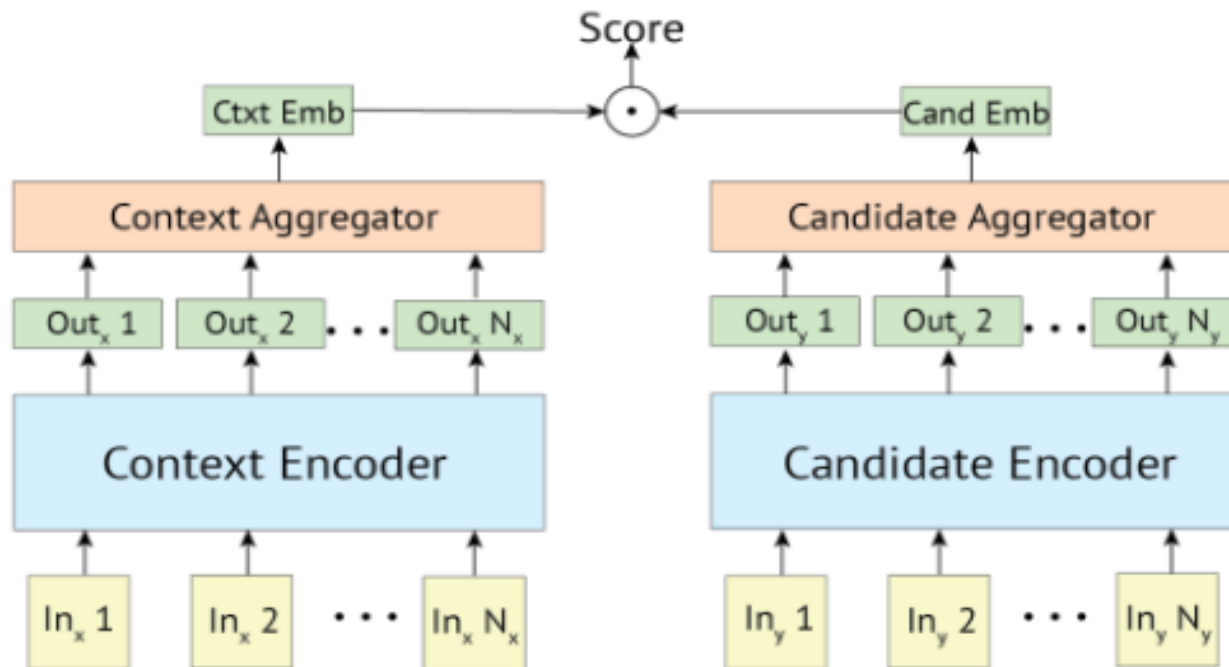
Models	Ubuntu			Douban						E-commerce		
	R ₁₀ @1	R ₁₀ @2	R ₁₀ @5	MAP	MRR	P@1	R ₁₀ @1	R ₁₀ @2	R ₁₀ @5	R ₁₀ @1	R ₁₀ @2	R ₁₀ @5
CNN (Kadlec, Schmid, and Kleindienst 2015)	0.549	0.684	0.896	0.417	0.440	0.226	0.121	0.252	0.647	0.328	0.515	0.792
LSTM (Kadlec, Schmid, and Kleindienst 2015)	0.638	0.784	0.949	0.485	0.537	0.320	0.187	0.343	0.720	0.365	0.536	0.828
BiLSTM (Kadlec, Schmid, and Kleindienst 2015)	0.630	0.780	0.944	0.479	0.514	0.313	0.184	0.330	0.716	0.365	0.536	0.825
MV-LSTM (Wan et al. 2016)	0.653	0.804	0.946	0.498	0.538	0.348	0.202	0.351	0.710	0.412	0.591	0.857
Match-LSTM(Wang and Jiang 2016)	0.653	0.799	0.944	0.500	0.537	0.345	0.202	0.348	0.720	0.410	0.590	0.858
Multi-View (Zhou et al. 2016)	0.662	0.801	0.951	0.505	0.543	0.342	0.202	0.350	0.729	0.421	0.601	0.861
DL2R (Yan, Song, and Wu 2016)	0.626	0.783	0.944	0.488	0.527	0.330	0.193	0.342	0.705	0.399	0.571	0.842
SMN (Wu et al. 2017)	0.726	0.847	0.961	0.529	0.569	0.397	0.233	0.396	0.724	0.453	0.654	0.886
DUA (Zhang et al. 2018)	0.752	0.868	0.962	0.551	0.599	0.421	0.243	0.421	0.780	0.501	0.700	0.921
DAM (Zhou et al. 2018)	0.767	0.874	0.969	0.550	0.601	0.427	0.254	0.410	0.757	0.526	0.727	0.933
IoI (Tao et al. 2019b)	0.796	0.894	0.974	0.573	0.621	0.444	0.269	0.451	0.786	0.563	0.768	0.950
MSN (Yuan et al. 2019)	0.800	0.899	0.978	0.587	0.632	0.470	0.295	0.452	0.788	0.606	0.770	0.937
BERT (Gu et al. 2020)	0.808	0.897	0.975	0.591	0.633	0.454	0.280	0.470	0.828	0.610	0.814	0.973
BERT-SS-DA (Lu et al. 2020)	0.813	0.901	0.977	0.602	0.643	0.458	0.280	0.491	0.843	0.648	0.843	0.980
SA-BERT (Gu et al. 2020)	0.855	0.928	0.983	0.619	0.659	0.496	0.313	0.481	0.847	0.704	0.879	0.985
BERT (ours)	0.820	0.906	0.978	0.597	0.634	0.448	0.279	0.489	0.823	0.641	0.824	0.973
ELECTRA	0.826	0.908	0.978	0.602	0.642	0.465	0.287	0.483	0.839	0.609	0.804	0.965
UMS _{BERT}	0.843	0.920	0.982	0.597	0.639	0.466	0.285	0.471	0.829	0.674	0.861	0.980
UMS _{ELECTRA}	0.854	0.929	0.984	0.608	0.650	0.472	0.291	0.488	0.845	0.648	0.831	0.974
BERT+	0.862	0.935	0.987	0.609	0.645	0.463	0.290	0.505	0.838	0.725	0.890	0.984
ELECTRA+	0.861	0.932	0.985	0.612	0.655	0.480	0.301	0.499	0.836	0.673	0.835	0.974
UMS _{BERT+}	0.875	0.942	0.988	0.625	0.664	0.499	0.318	0.482	0.858	0.762	0.905	0.986
UMS _{ELECTRA+}	0.875	0.941	0.988	0.623	0.663	0.492	0.307	0.501	0.851	0.707	0.853	0.974

3.4 프로젝트 소개

Bi-Encoder Model

입력 데이터와 추론 결과를 각각의 Encoding을 통해 학습 하는 방법

- 자연어 텍스트 데이터를 다중 차원의 Vector 값으로 임베딩
- 학습시에 입력-결과를 같이 학습시켜 서로 비슷한 Vector값으로 임베딩 되도록 함
- 정답 입력은 '입력-결과'의 임베딩이 가깝도록, 오답 입력은 '입력-결과'의 임베딩이 멀어지도록 학습



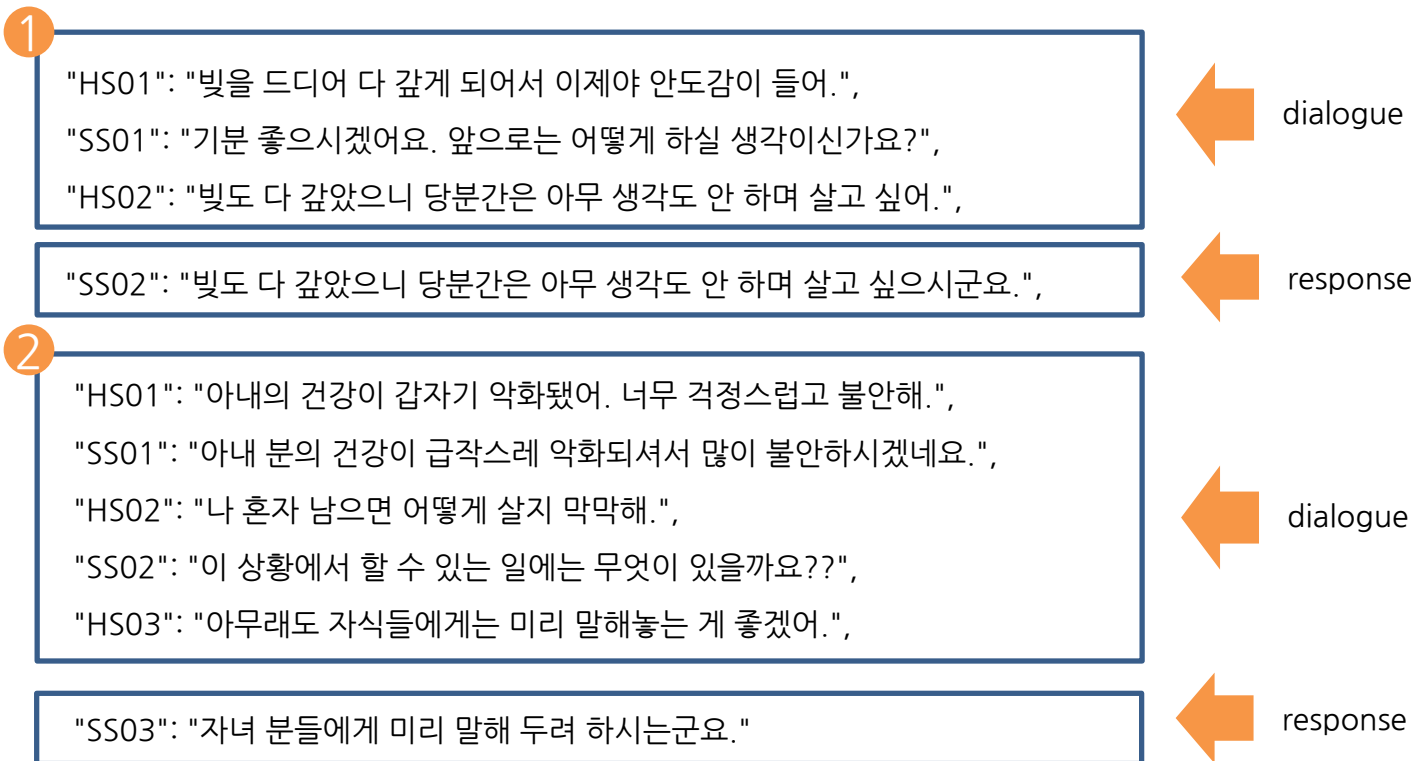
Bi-Encoder 구조도

3.4 프로젝트 소개

(Multi-turn)Response Selection

문맥이 있는 대화가 주어졌을 때, 사전에 정해진 응답 후보 중에서 **적절한** 응답을 선택하는 태스크

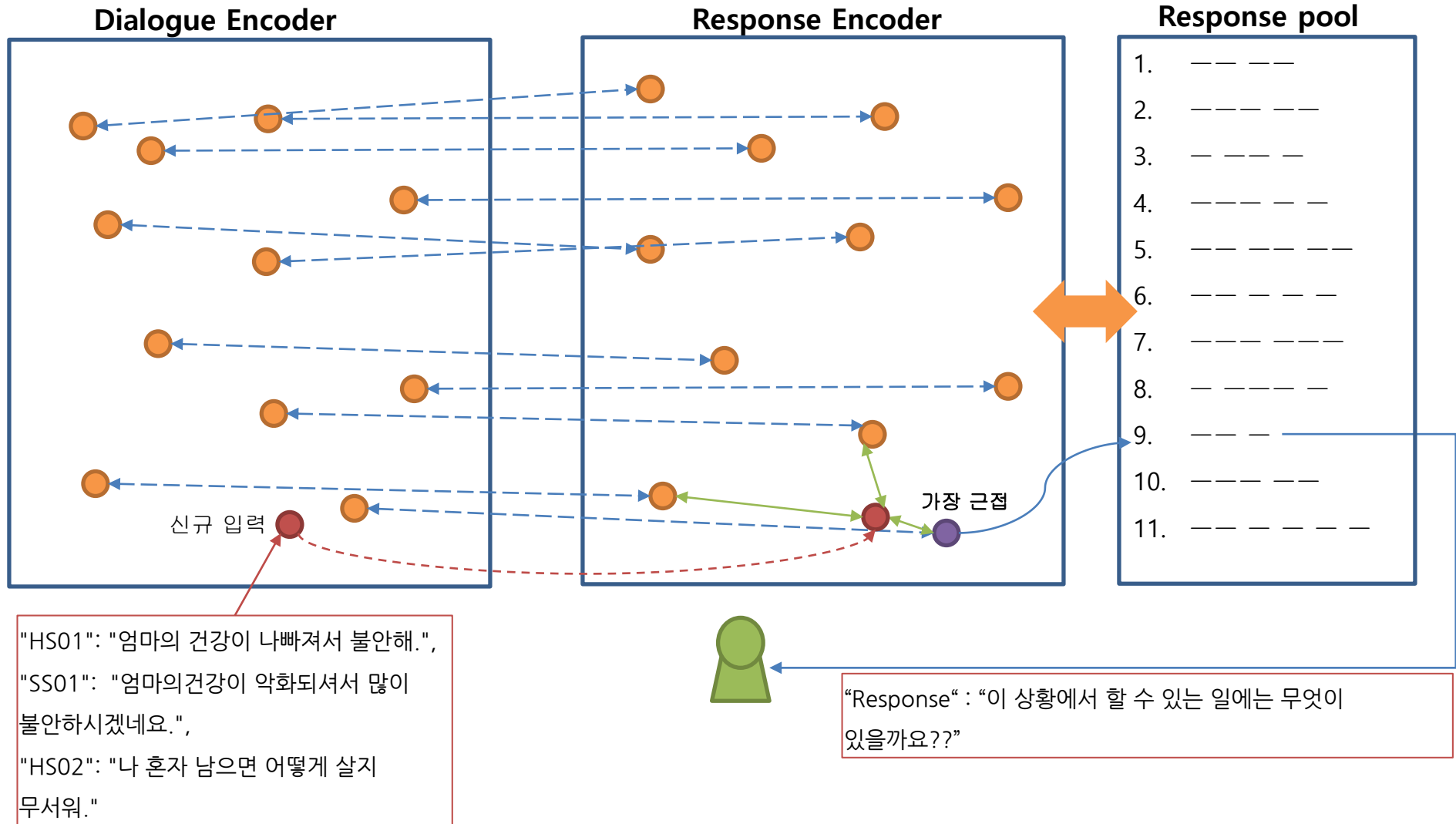
- 모델 학습에 dialogue, response으로 구성



출처) AI Hub 감성 대화 말뭉치
<https://aihub.or.kr/aidata/7978>

3.4 프로젝트 소개

Response Selection 도식화



3.4 프로젝트 소개

커리큘럼

1주차

- (이론)언어 모델 이해
- (이론)다운스트림 태스크 개요
- (이론)Response Selection 개념
- (실습)학습 데이터 분석
- (리포트)학습 데이터 분석 결과

2주차

- (실습)학습 데이터 구축
- (이론)Response Selection 학습 가이드
- (실습)Response Selection Fine-tuning 수행
- (리포트)학습 진행/결과

4주차

- (이론)다양한 다운스트림 태스크 소개
- (실습)다운스트림 태스크 서비스 분석
- (실습)Response Selection 품질 고도화
- (리포트)프로젝트 최종 결과/보고서

3주차

- (이론)Faith Search 개념 이해
- (실습)Inference 및 유사도 도출
- (리포트)Response Selection 품질 고도화 방법
- (실습)학습 데이터 보강



Thank you



The easy way to start customer service chatbot
and application strategy of it for your business.