

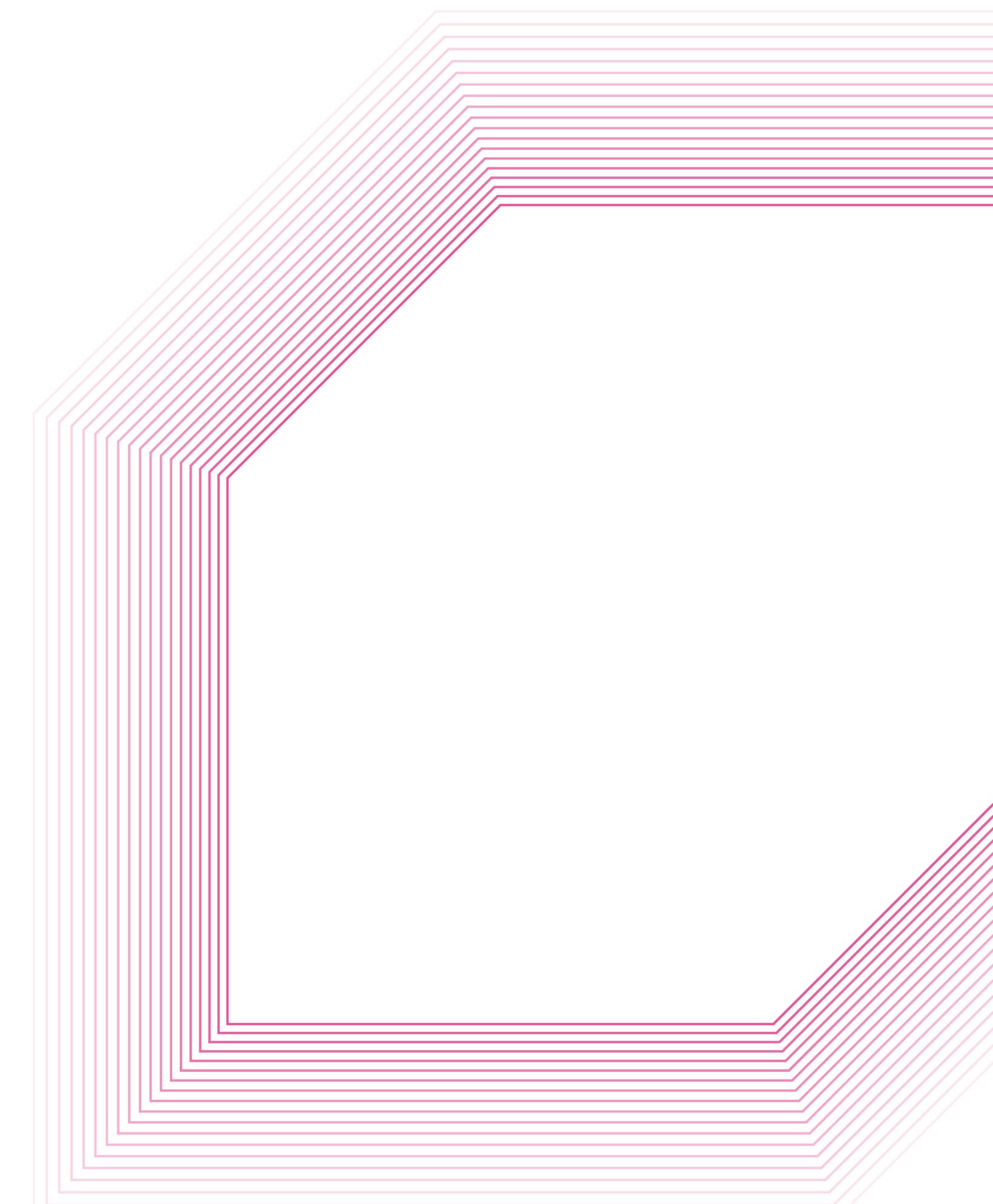
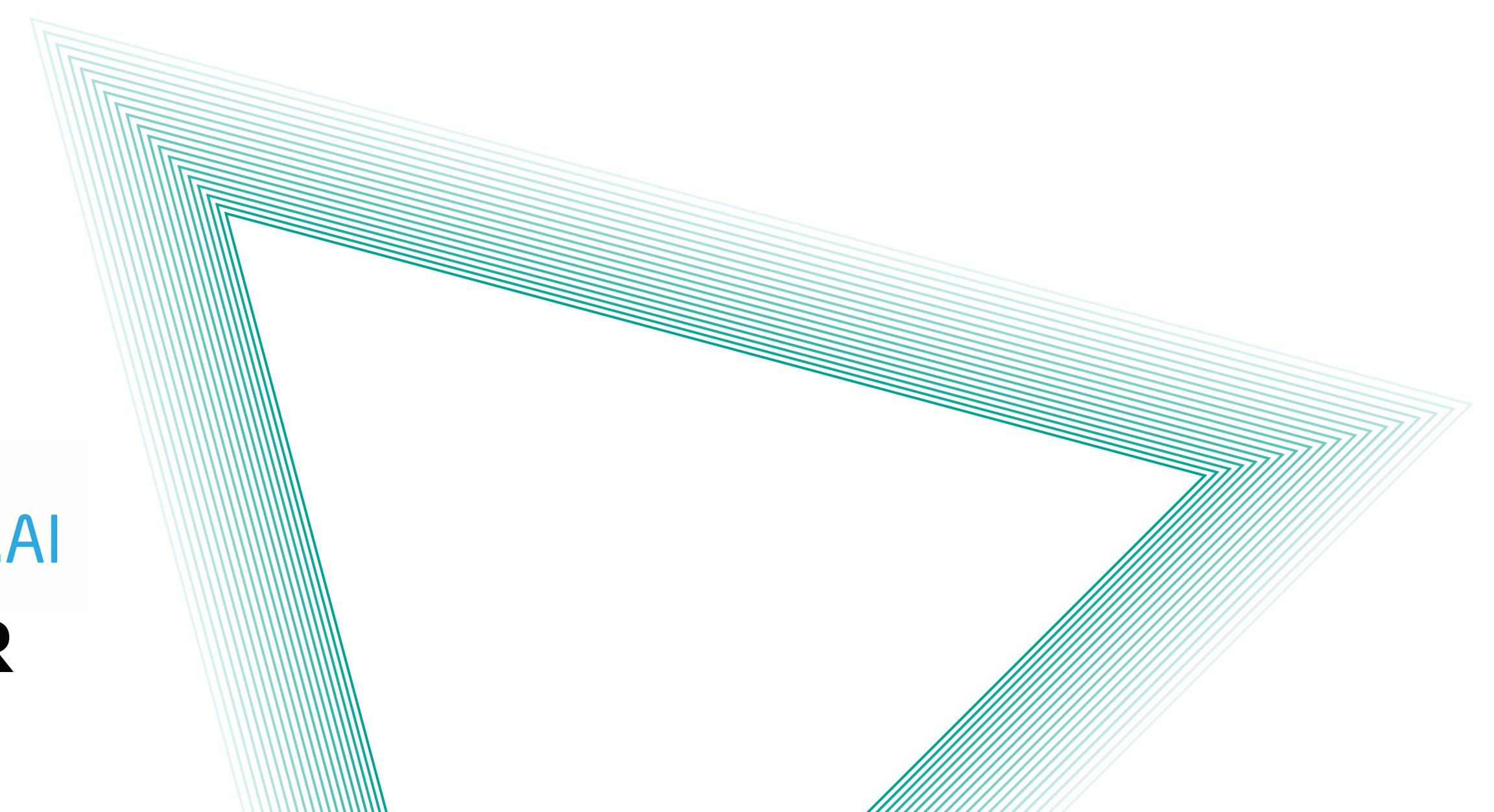
DEVIEW
2018

인공지능이 인공지능 챗봇을 만든다.

이재원

Company.AI

NAVER



CONTENTS

DEVIEW
2018

1. CHATBOT?? CHATBOT!!
2. 스스로 언어를 배우는 인공지능 챗봇 만들기
3. 말뭉치 줄게, 챗봇 다오!!
4. 인공지능 챗봇의 말하기 실력은?
5. 클러스터링, 그리고 파라미터 튜닝
6. 상용 서비스 현황 및 모델 성능

1.
CHATBOT??
CHATBOT!!

1.1 응? 챗봇?

DEVIEW
2018

잘 못 알아 들었습니다

다시 한 번 말씀해
주시겠습니까?



Chat하기 힘든 Chatbot

1.2 왜 챗봇인가? (1/2)

DEVIEW
2018

국내뿐만 아니라 전세계적으로 챗봇 시장 규모 기하급수적으로 증가

1. 생활 속 편리함, 모바일 음성검색, 인건비 등 비용절감 등의 목적으로 챗봇 사용 니즈 증가
2. 국내에서도 챗봇은 금융, 커머스, 유통, 기업 관리 및 공공기관 등 다방면에 걸쳐 영역 확장 중

한국경제

스타트업 'AI 챗봇' 개발 경쟁... 3조 시장 불붙었다

동아일보

"홈쇼핑 물건 살때도, 숙소 찾을때도... 이젠 앱 대신 챗봇"

금융권 챗봇 돌풍, 내년까지 도입 금융회사 47개사로 확대

아이뉴스24

KT, 챗봇 '전대리'에게 전표처리 맡긴다

news 1
KOREA

정부, 성범죄 피해자 돕는 상담챗봇 개발한다

1.2 왜 챗봇인가? (2/2)

DEVIEW
2018

국내뿐만 아니라 전세계적으로 챗봇 시장 규모 기하급수적으로 증가

1. 글로벌 기업 및 국내 대기업 챗봇 관련 스타트업 적극 인수합병으로 경쟁 우위 달성 노력



16년 9월, API.AP 인수 → “Dialogflow”로 서비스

Google Assistant 내부 적용 빌더 사용

국내 챗봇 스타트업 “플런티” 인수(삼성이 국내 스타트업 인수한 첫 사례)

→ 기존 “빅스비”와 접목된 서비스 개선에 활용될 것으로 예상

AI연구 전문 자회사 “카카오 브레인” 설립

→ 인공지능 스피커보다 **챗봇 사업의 우선순위를 높여** 많은 투자 및 연구 진행 중

챗봇 사업을 위해 여러 회사와 합병 실시

→ 16년 6월 “완드 랩스”, 17년 1월 “말루바”, 18년 5월 “시맨틱 머신즈” 인수

지능형 챗봇 개발 사내벤처 “단비” 육성

→ LG U+ 고객센터 적용 후 단순 문의 해결 10배 증가 효과

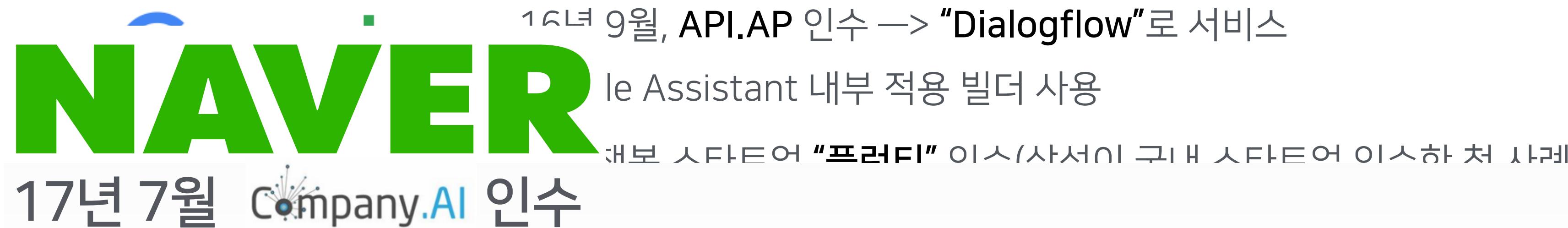


1.2 왜 챗봇인가? (2/2)

DEVIEW
2018

국내뿐만 아니라 전세계적으로 챗봇 시장 규모 기하급수적으로 증가

1. 글로벌 기업 및 국내 대기업 챗봇 관련 스타트업 적극 인수합병으로 경쟁 우위 달성 노력



→ 네이버와 라인이 공동 개발 중인 인공지능 플랫폼 Clova에 탑재될 핵심 기술 녹여내는 중



→ 인공지능 스피커보다 챗봇 사업의 우선순위를 높여 많은 투자 및 연구 진행 중

챗봇 사업을 위해 여러 회사와 합병 실시

→ 16년 6월 “완드 랩스”, 17년 1월 “말루바”, 18년 5월 “시맨틱 머신즈” 인수

지능형 챗봇 개발 사내벤처 “단비” 육성

→ LG U+ 고객센터 적용 후 단순 문의 해결 10배 증가 효과

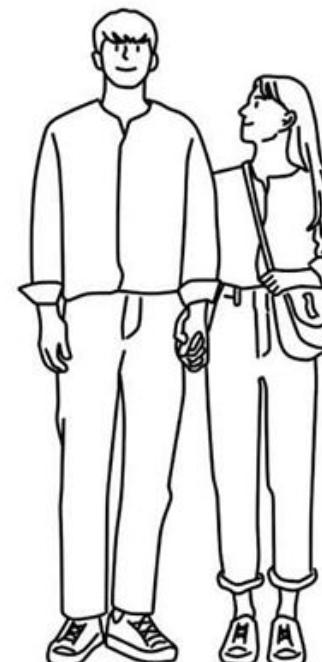


2. 스스로 언어를 배우는 인공지능 챗봇 만들기

2.1 사람이 언어를 배우는 과정

특정 상황 속에서
눈, 귀를 이용하여 말을 들음

실생활 중 특정상황에서
배운 말을 적용



두뇌를 이용하여
해당 상황 분석

특정 상황 속에서
말의 의미 파악

2.2 사람에게 새로운 언어가 어려운 이유

각 언어마다 문법과 구조가 달라, 일반화 어려움

1. 언어마다 주어, 동사 등 문장의 어순이 다름
2. '卜' 다르고, 'ㅏ' 다른 한국어
3. 'a', 'an', 'the' 등 관사의 역할이 중요한 영어

중국어 영어 스페인어 <-> 일어 한국어 등

2.3 그렇다면 인공지능에게 언어 학습이란?

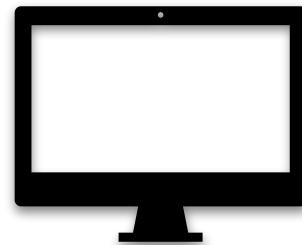
DEVIEW
2018

말뭉치 데이터 중 주변 단어를 보고 의미를 파악하여 feature 추출

단어의 주변을 보면 그 단어를 안다.

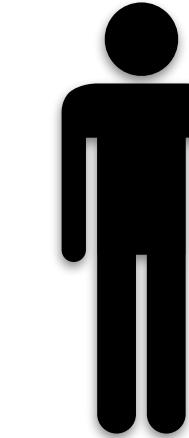


단어 의 주변 을 보면 그 단어 를 안다



단어: ??
주변: ??
보면: ??

단어[명사]: 분리
하여 자립적으로
쓸 수 있는 말



Text

문장의 각 단어의 고유 의미보다 주변 단어와 연관지어 단어(토큰)별 의미 부여

2.4 인공지능이 언어를 배우는 과정 (1/3)

DEVIEW
2018

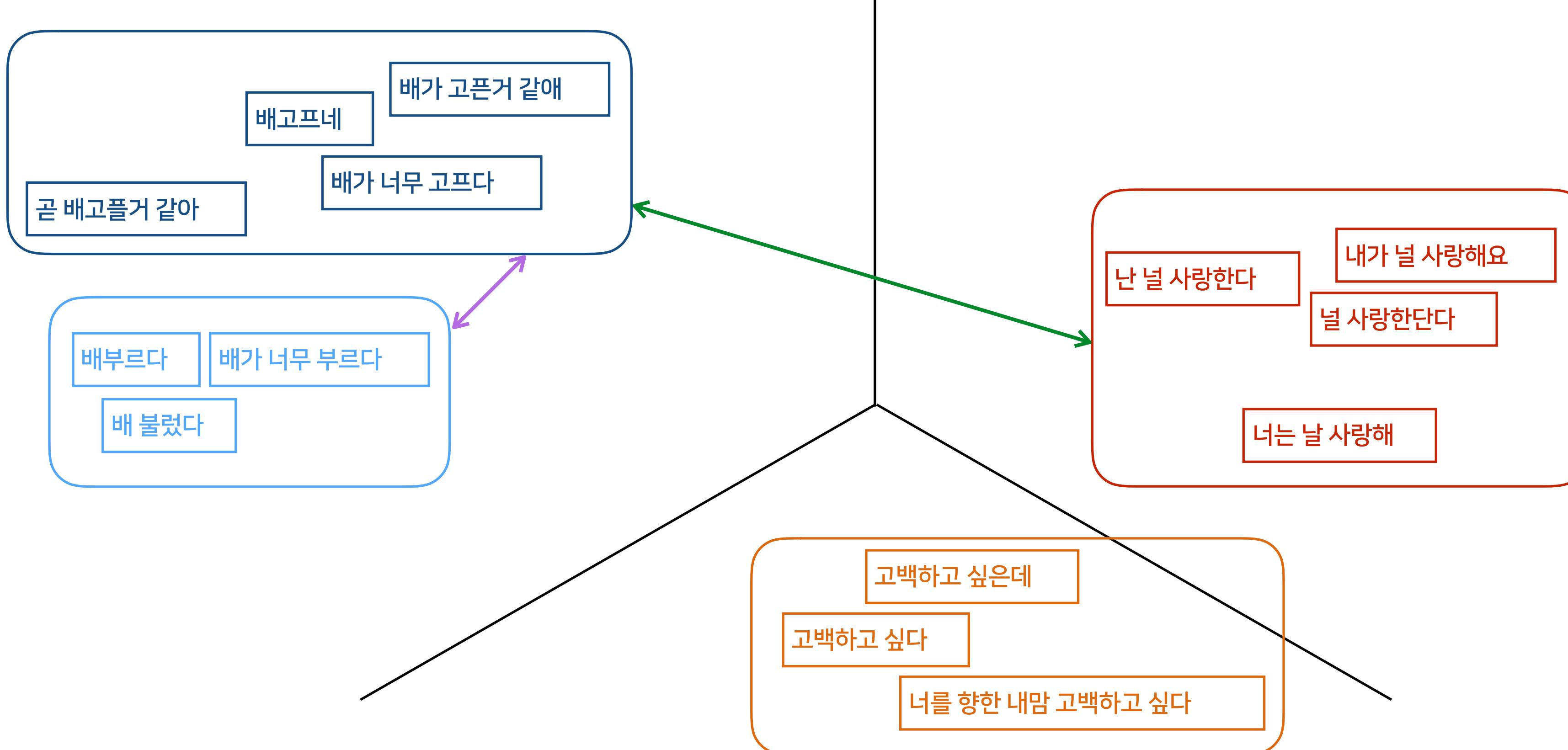
Vector Representation

1. 주어진 Vector Space에 각 문장의 Embedding Vector를 적절하게 배치
2. 각 문장별 Embedding Vector가 **가까울수록** 같은 의미, **멀수록** 다른 의미
3. 각 n차원의 공간에서 n개의 원소는 문장의 특징을 표현한다
4. 각 도메인별 말뭉치로 학습된 N-hot Embedding Vector
5. Stanford의 Glove, Facebook의 fastText, NAVER 내부용 Global Embedding Vector

2.4 인공지능이 언어를 배우는 과정 (2/3)

DEVIEW
2018

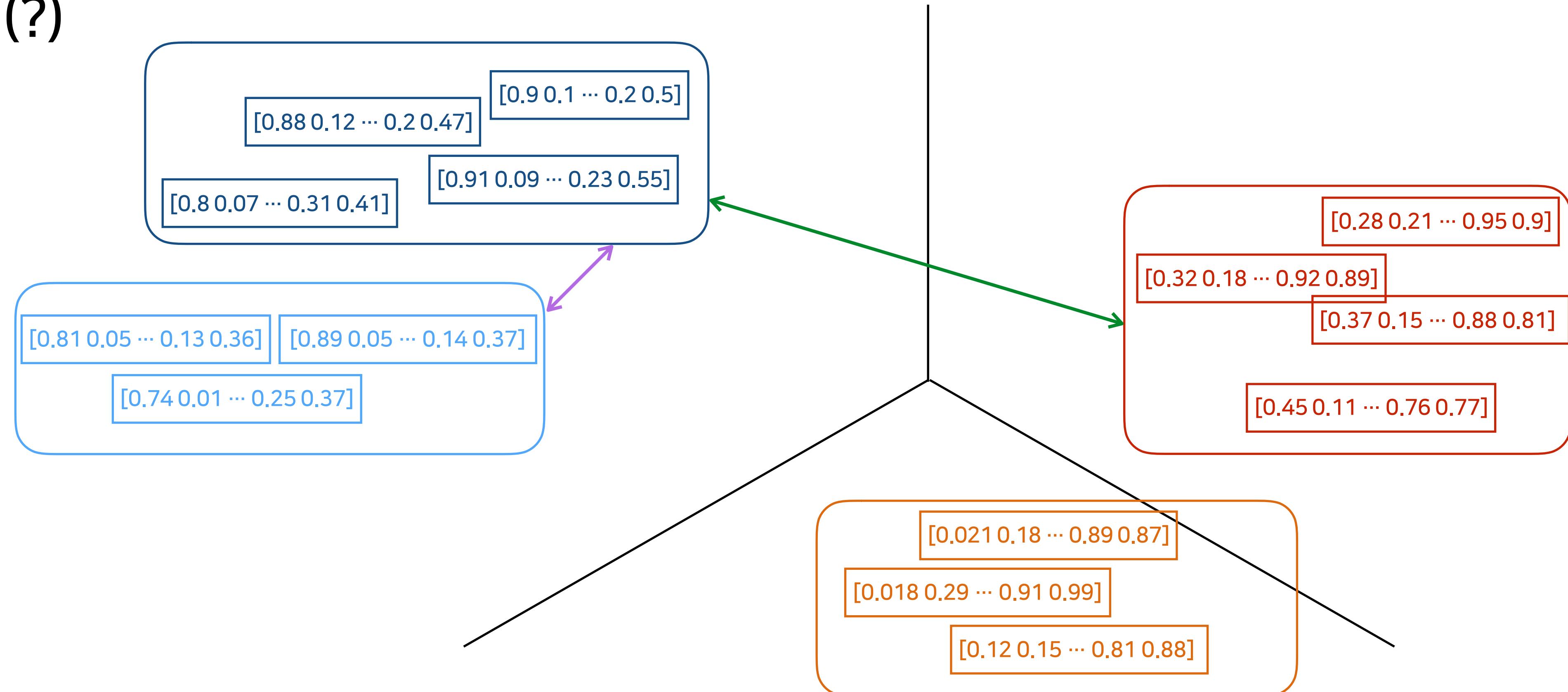
인간의 노



2.4 인공지능이 언어를 배우는 과정 (3/3)

DEVIEW
2018

인공지능의 뇌(?)



2.5 인공지능 스스로 배울 수 있지 않을까?

DEVIEW
2018

어차피 내가 만드는 숫자가
embedding vector인데,
계속 배우다보면 정확해지지 않을까?

아무도 도와주지 않아도
할 수 있을 것 같아요!



2.6 AutoML 이란? (1/2)

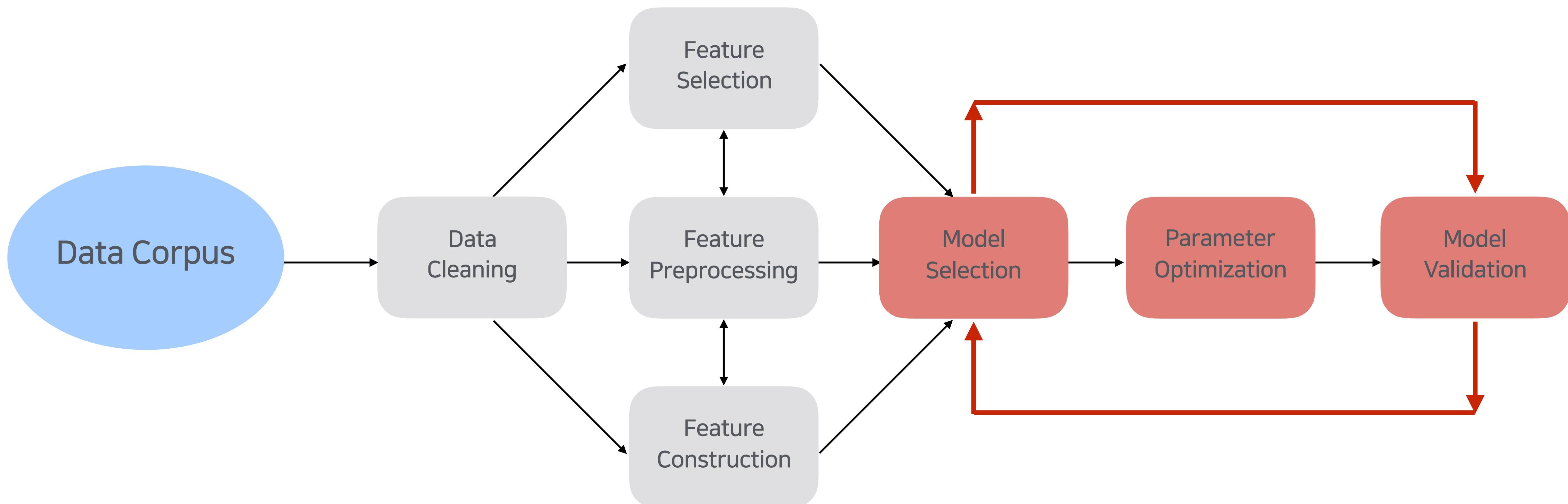
DEVIEW
2018

진화된 인공지능 프로그램

1. 자신의 환경에 대해 학습하는 복잡한 코드와 알고리즘을 복합적으로 생성하는 신경망 프로세스
2. 컴퓨터 언어를 몰라도 프로그램 설계 가능
3. 부모가 되는 인공지능 기술과 자식 개념의 머신러닝 프로그램으로 수천번 반복
4. 주어진 몇가지 학습 키워드로 스스로 프로그램 설계 및 제공

2.6 AutoML 이란? (2/2)

그렇다면 우리가 사용할 AutoML은?

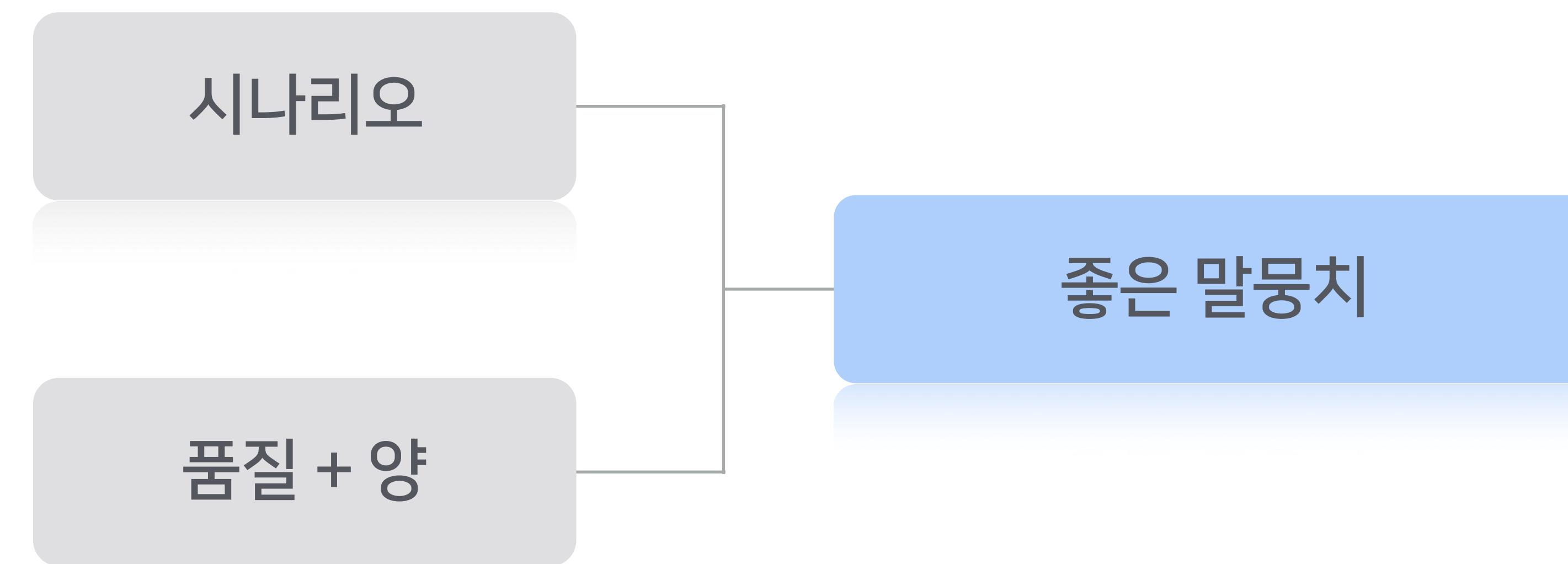


3. 말뭉치 줄게, 챗봇 다오!!

3.1 우리에게 필요한건 말뭉치뿐!

좋은 말뭉치는 좋은 선생님이다.

1. 말뭉치는 상황별 다양한 질문과 적절한 답변을 맵핑한 시나리오
2. 말뭉치는 인공지능이 배울 언어의 영역이자, 한계
3. 말뭉치는 챗봇의 품질을 결정

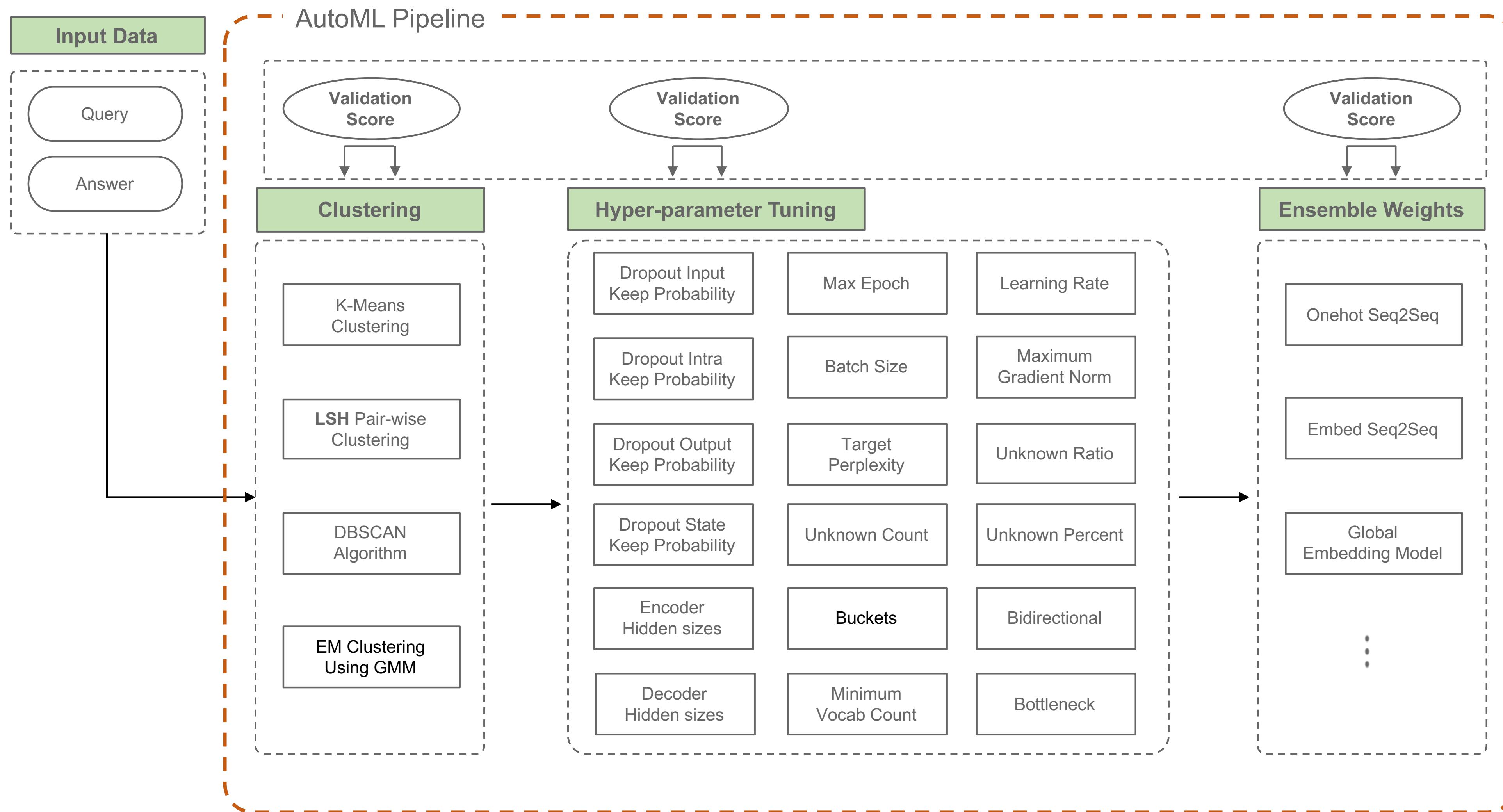


3.2 챗봇의 변신은 무죄! (1/2)

챗봇은 주어진 말뭉치로 학습되어 AutoML을 통해 끊임없이 변화, 발전

1. 스스로 학습하면서 평가 및 수정, 그리고 다시 학습
2. 알고리즘 선택, 하이퍼 파라미터 튜닝, 반복학습, 모델평가 모두 자동화
3. One-click으로 수준 높은 모델 구축 가능
4. 직관에 의존하는 모델링 개선 가능
5. 범용성 높은 모델로 서로 다른 도메인 및 다른 데이터 형태에도 쉽고 다양하게 적용 가능

3.2 AutoML과 함께라면! (2/2)



3.3 자연어처리, 챗봇, 그리고 AutoML

DEVIEW
2018

하지만, 아직...

1. 말뭉치만으로 인공지능 스스로 언어의 특징을 파악하고 모델을 만들기에는 역부족
2. Q-A 말뭉치를 학습에 필요한 시나리오로 변환하기 위한 리소스 부족
3. 도메인 키워드 만으로 아무 도움없이 인공지능 스스로 챗봇을 만들기에는 역부족
4. 각 도메인에 적합한 모델별 파라미터 조합에 대한 확신 및 검증 부족

3.3 그렇다면 AutoML 적용할 수 없나?

“방대한 말뭉치 효율적 관리 방법” 그리고 “모델 검증 및 테스트 구조”만 있다면!

Clustering

말뭉치 속 여러 문장을 모델 스스로 의미 구분하여
효율적으로 시나리오 관리

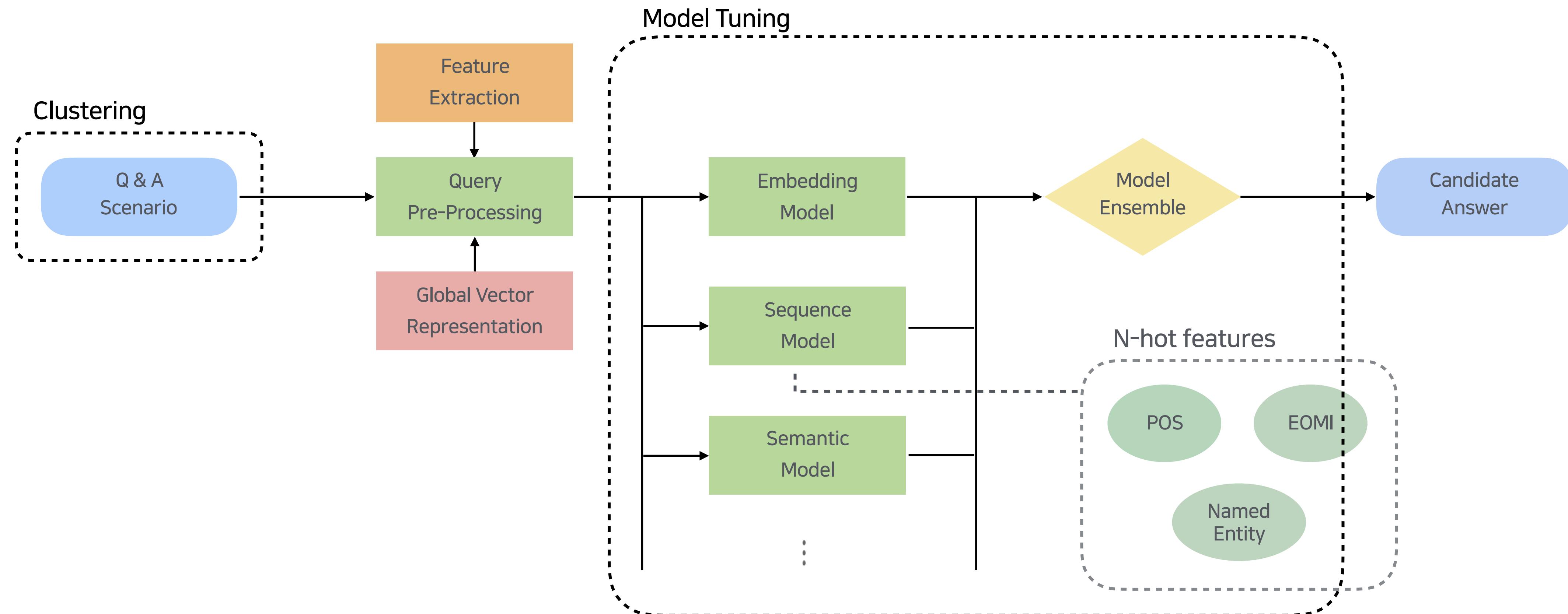
도메인별 특징을 반영하여
최적의 하이퍼 파라미터 조합으로 생성된 모델

Hyper-
parameter Tuning

3.4 AutoML을 도와줄 Framework (1/7)

DEVIEW
2018

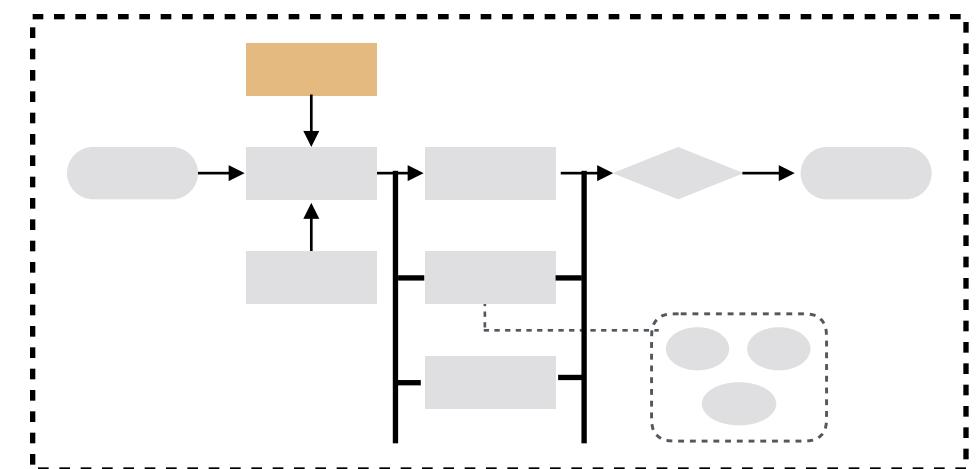
쉽고 빠른 Company.AI Chatbot Builder Framework



3.4 Chatbot Module (2/7)

Feature Extraction: 언어특징 추출 라이브러리(rama)

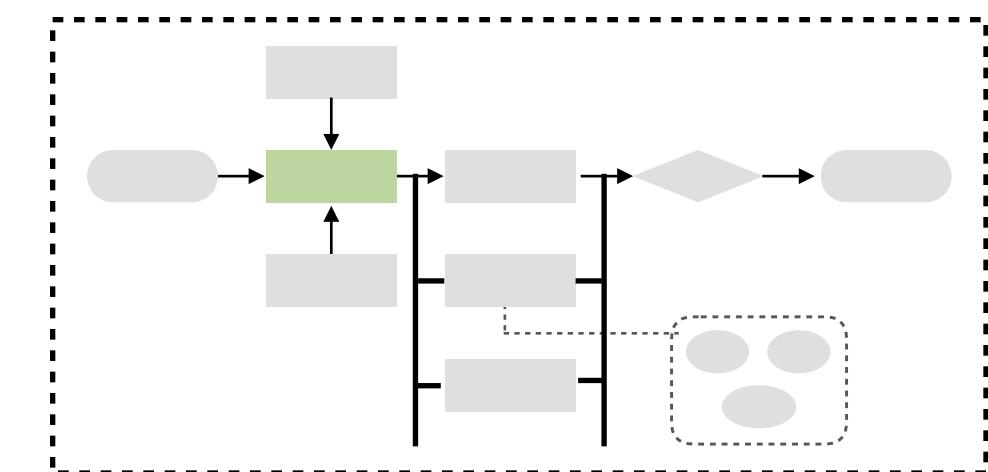
1. 형태소 분석
 1. 다국어 지원, 분산 처리 기반 분석 지원
2. 입력 문장 및 개체 분석 처리
 1. 학습형 / 사전형 분석
 2. 네이버 개체명 체계 지원 / 사용자 정의 지원
 3. 형태소 및 유니코드 정규화
3. 사전 관리
 1. 형태소 후처리 사전 관리
 2. 개체 사전 관리



3.4 Chatbot Module (3/7)

Query Pre-processing: 쿼리 전처리 및 학습 컨트롤러(baloo)

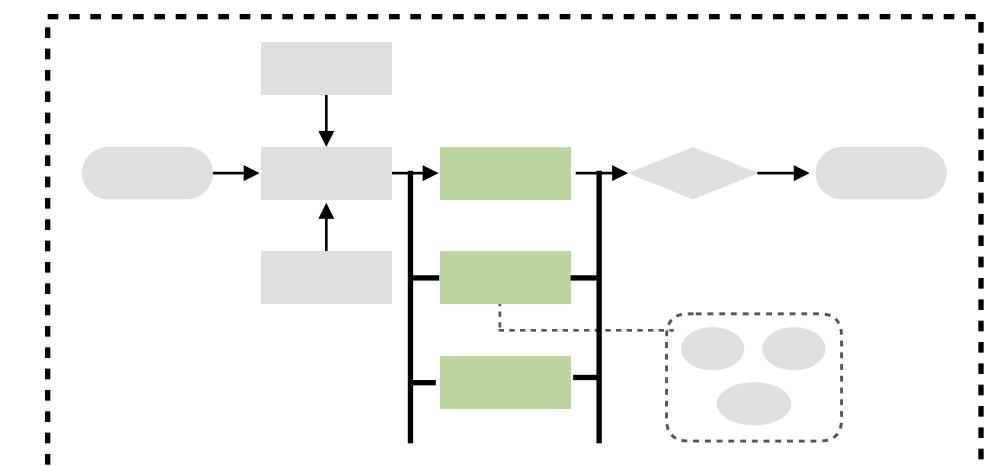
1. 학습에 필요한 모든 쿼리 및 인입된 신규 쿼리 전처리
 1. 문장 부호, 특수문자, 숫자, 외래어 등 전처리
 2. 말뭉치의 Vocab 파일, 문장별 Tokenized된 파일 등 지정된 HDFS 경로에 저장
2. 전체 모델 학습 컨트롤러
 1. NAVER GPU([c3dl](#)) 학습 요청 및 현황 관리
 2. 학습 완료 후 서비스에 필요한 모델별 tensorflow pb 파일 등 생성
 3. 학습 파라미터와 비학습 파라미터 별도 관리



3.4 Chatbot Module (4/7)

Predict Answers with Dialogue Model (akella)

1. 한국어 및 대화형 데이터에 최적인 Seq2Seq 기반 Tensorflow Model
2. 형태소 분석된 **질의 시퀀스**로 **답변 시퀀스**를 생성



Cell

Architecture

Attention

Long Short-Term Memory

N-hot Representation

Multiplicative Attention

Gated Recurrent Unit

Bidirectional Encoder

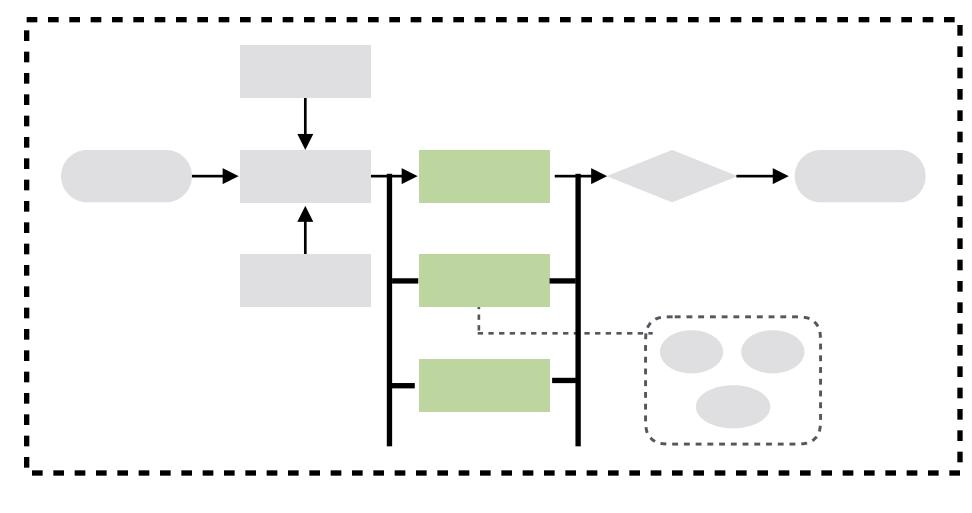
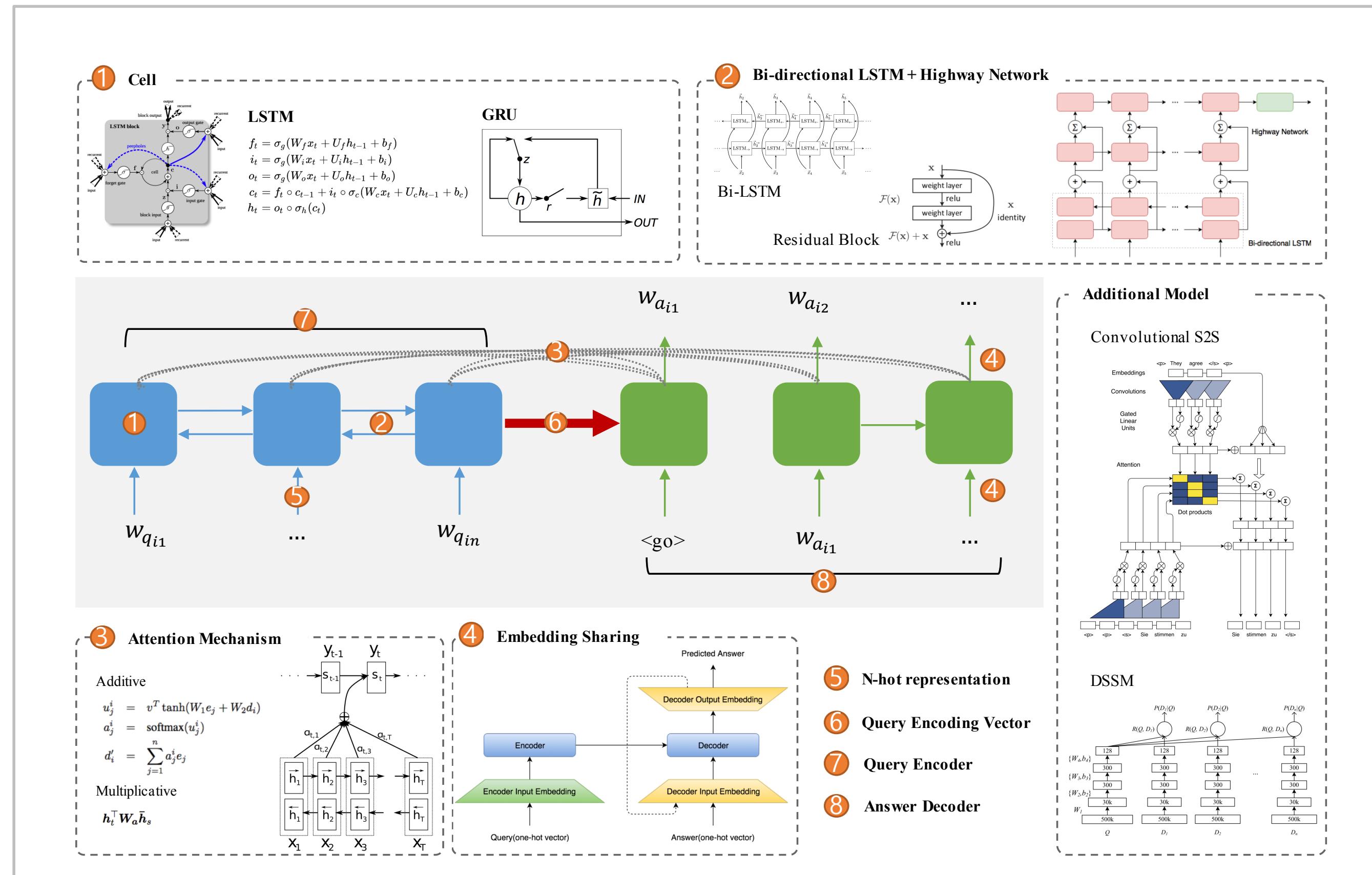
Additive Attention

Embedding Sharing

Highway Network

3.4 Chatbot Module (5/7)

“사람이 대화를 배워가는 과정과 유사한 형태로...”



다양한 문장과 문법을 공부하며:

Vector Representation / N-hot Representation

중요한 정보를 기억하며: LSTM

추가로 전후 문맥을 보고: Bi-directional

추가로 반복학습하며: Highway Network

정답 위주로 Feedback을 주며: Attention

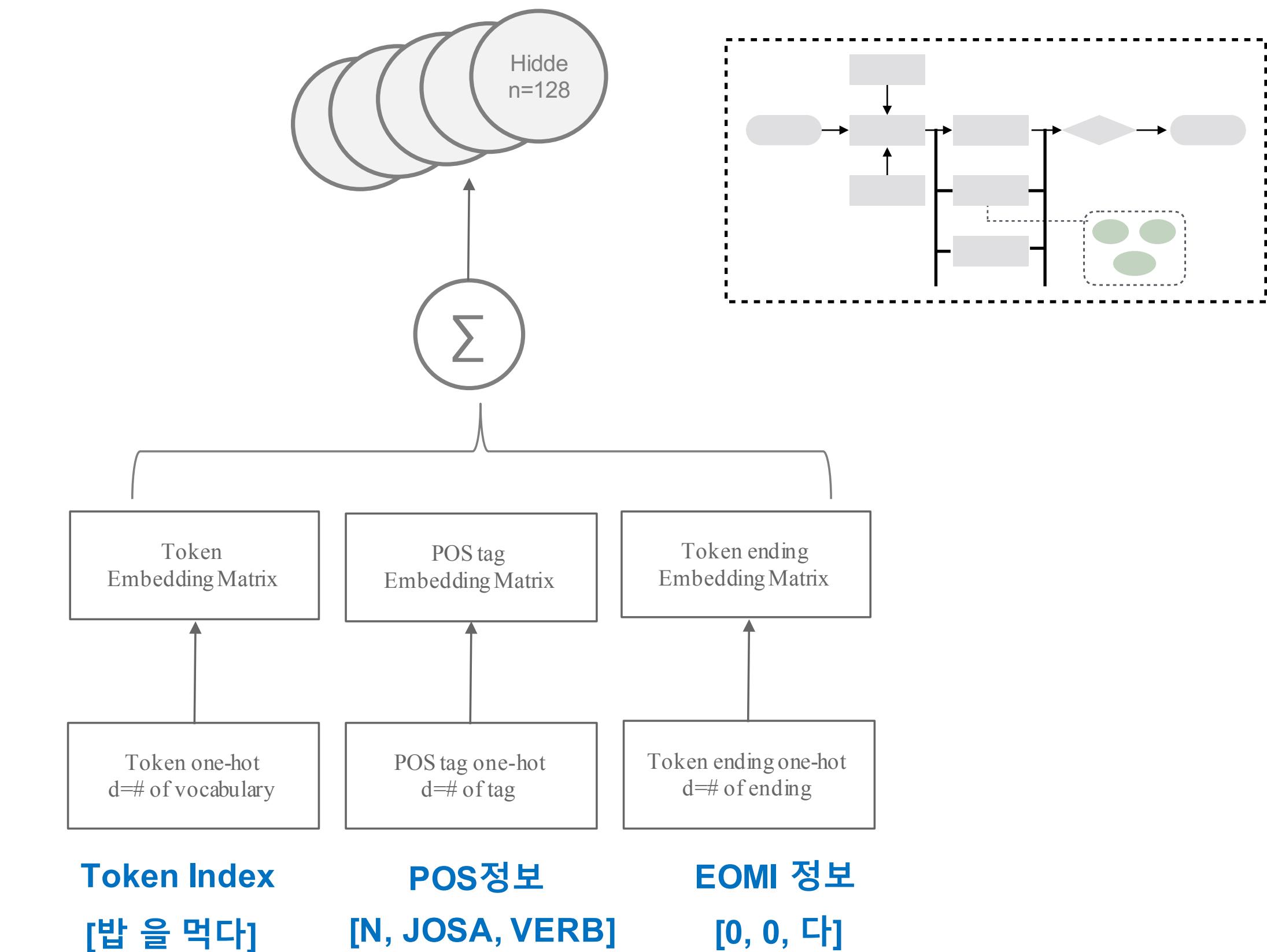
다양한 모델의 다수결로 답변을 추출:

Model Ensemble

3.4 Chatbot Module (6/7)

N-hot Input Model

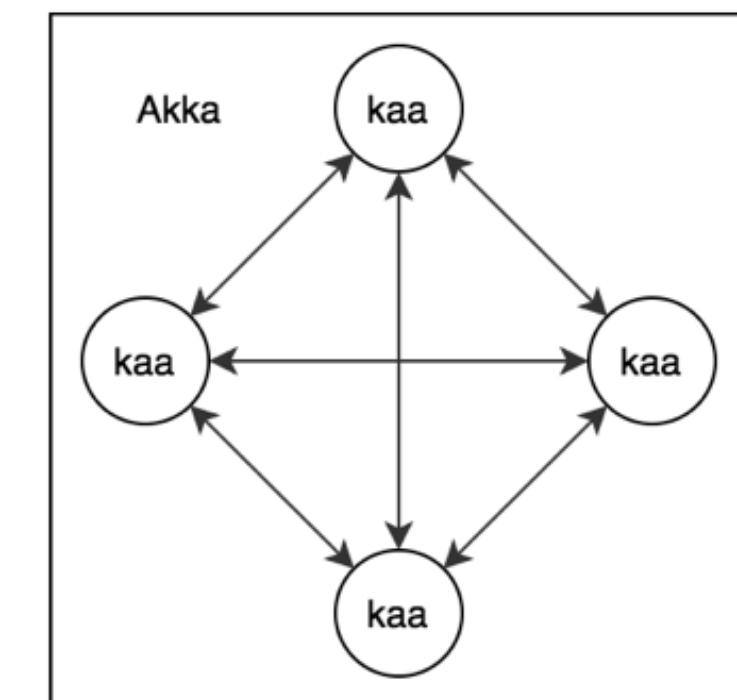
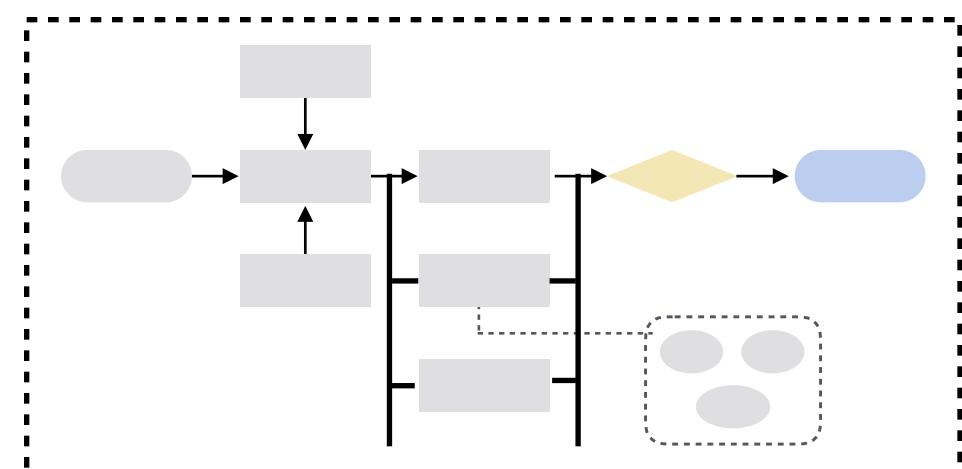
- A. Token 정보 이외 문장 이해에 도움을 주는 다른 정보를 같이 활용하는 방안
- B. 서술어의 다양한 변형을 적절하게 처리할 수 있는 방법으로 원형을 token으로 활용하면 효과적
- C. 이 때 어미 변형이 담고있는 정보를 손실시키지 않기 위하여 어미만 별도의 feature로 활용



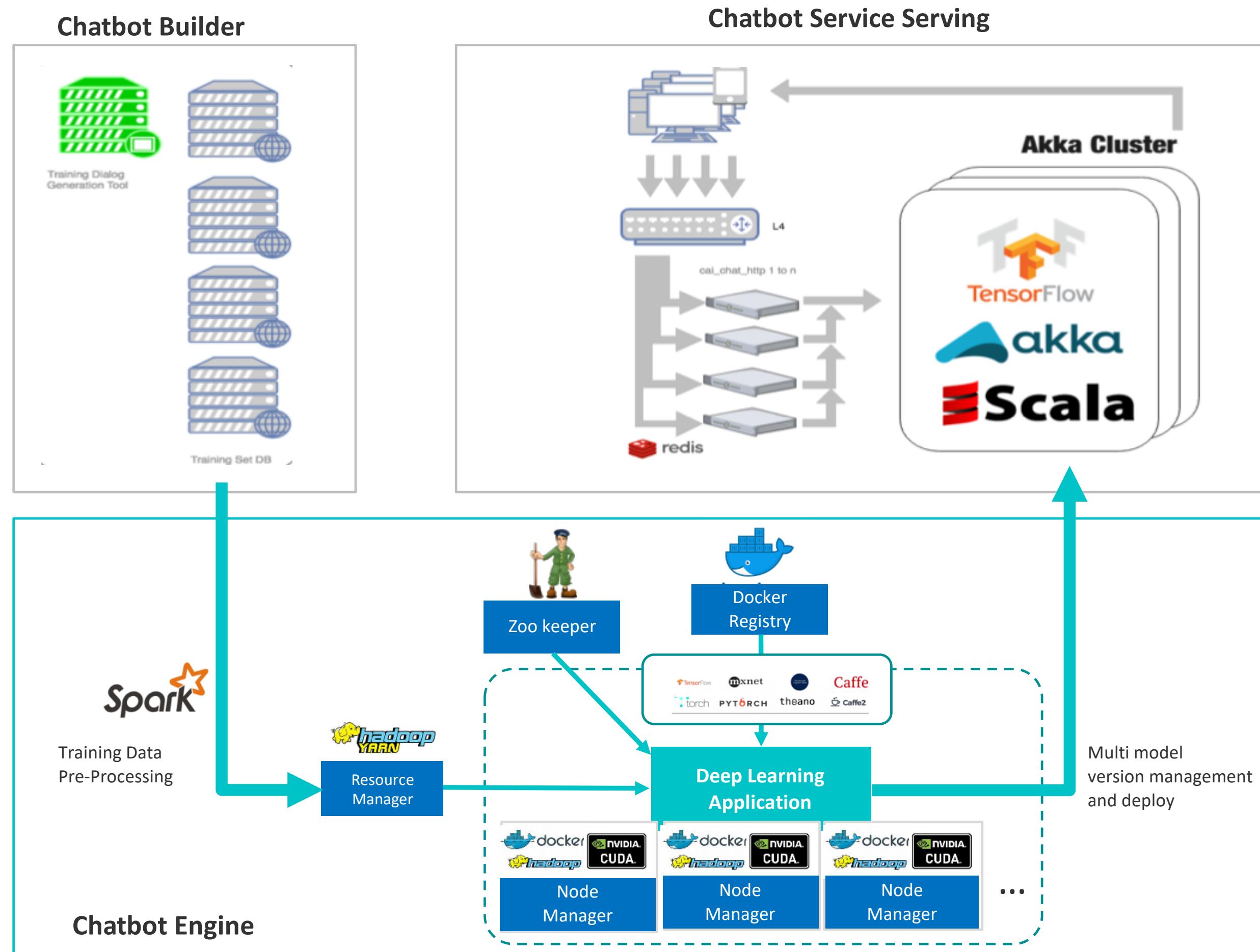
3.4 Chatbot Module (7/7)

Serve to multi-platform (kaa)

1. 해당 쿼리에 대한 여러 모델의 답변을 **양상블**하여 답변 제공
2. 다양한 경로로 인입되는 쿼리를 실시간 처리 후, 플랫폼으로 **답변 서비스**
3. 다중화 및 샤딩
 - A. 무중단 서비스를 위해 Akka를 사용하여 여러대 서버의 모듈간 통신
 - B. 서버자원 절약을 위해 사용량에 따른 서버 사용량 차등 적용



3.5 Chatbot Builder Framework



그리하여 최종 Process는?

Build → Train → Tune → Deploy

1단계:

- 사용자 대화 시나리오를 수집 및 정제 [Hadoop]

2단계:

- [GPU] 사용하여 말뭉치를 학습할 수 있도록 구성
- [Tensorflow 및 다양한 ML Framework] 활용

3단계:

- 반복적으로 모델 튜닝 결과 서버에 배포[Server]

4. 인공지능 챗봇의 말하기 실력은?

4.1 Chatbot 성능평가 기준 (1/2)

DEVIEW
2018

성능 좋은 Chatbot 이란?

부적절한 답변율

친근함

정확한 개념 이해

견고한 모델

사용하기 쉬운 챗봇

재사용 여부

명확한 답변

자연스러운 답변

4.1 Chatbot 성능평가 기준 (2/2)

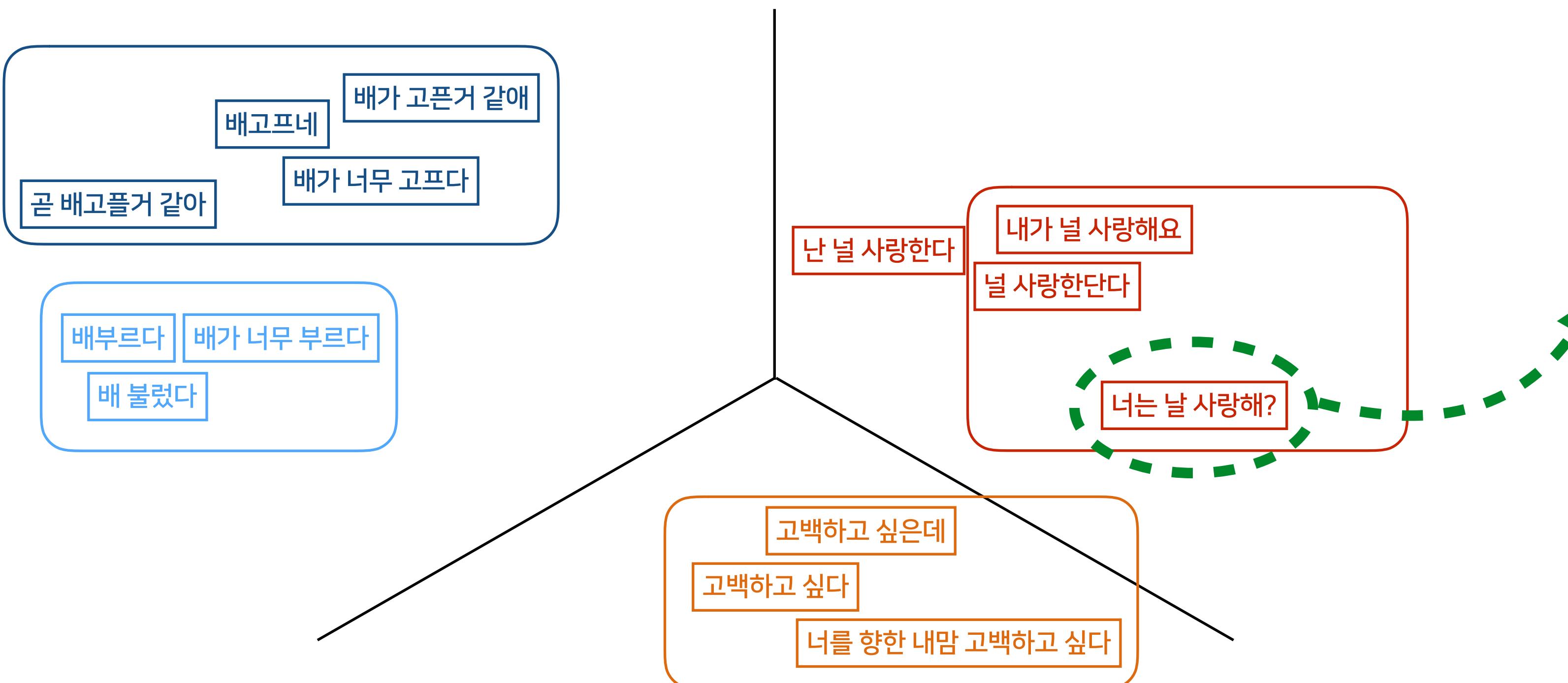
Clova 한국어대화 4점 척도 모델 평가 예시



4.2 Chatbot 자동평가 방법

모델 스스로 각각의 쿼리를 Embedding Vector Space

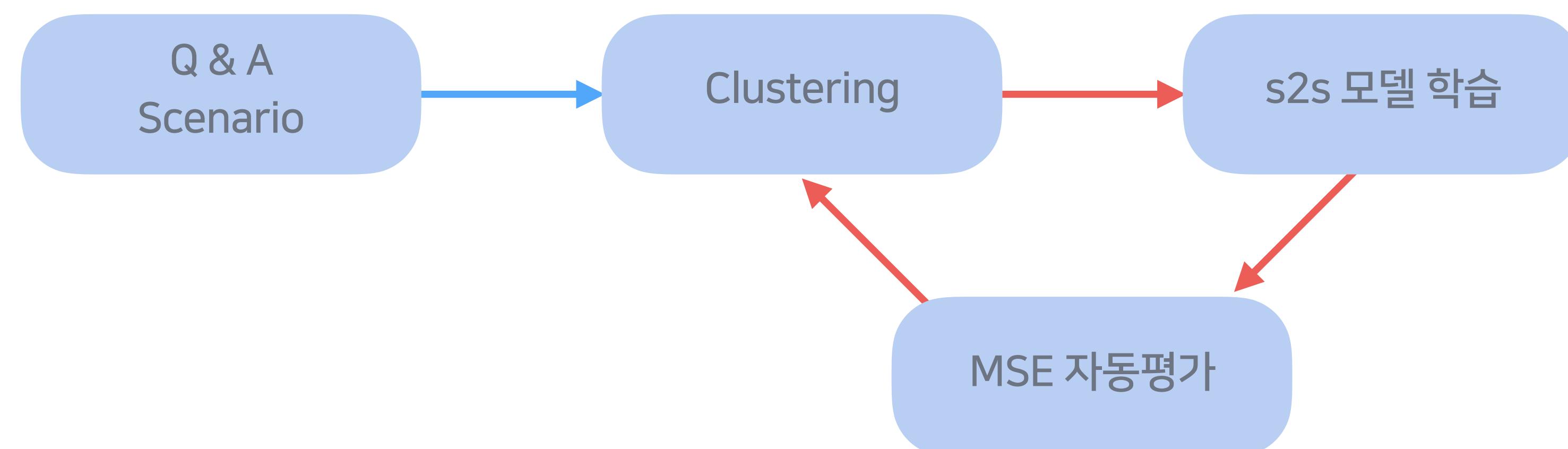
1. 시나리오에 포함된 쿼리 중 1~2개를 제외하고, 그 쿼리는 오직 평가에만 사용
2. 해당 평가 쿼리는 정답과 오답 시나리오 양쪽 모두 자동평가 진행



4.2 Clustering 자동평가 방법

군집간 분산 **최대화** 및 군집내 분산 **최소화**

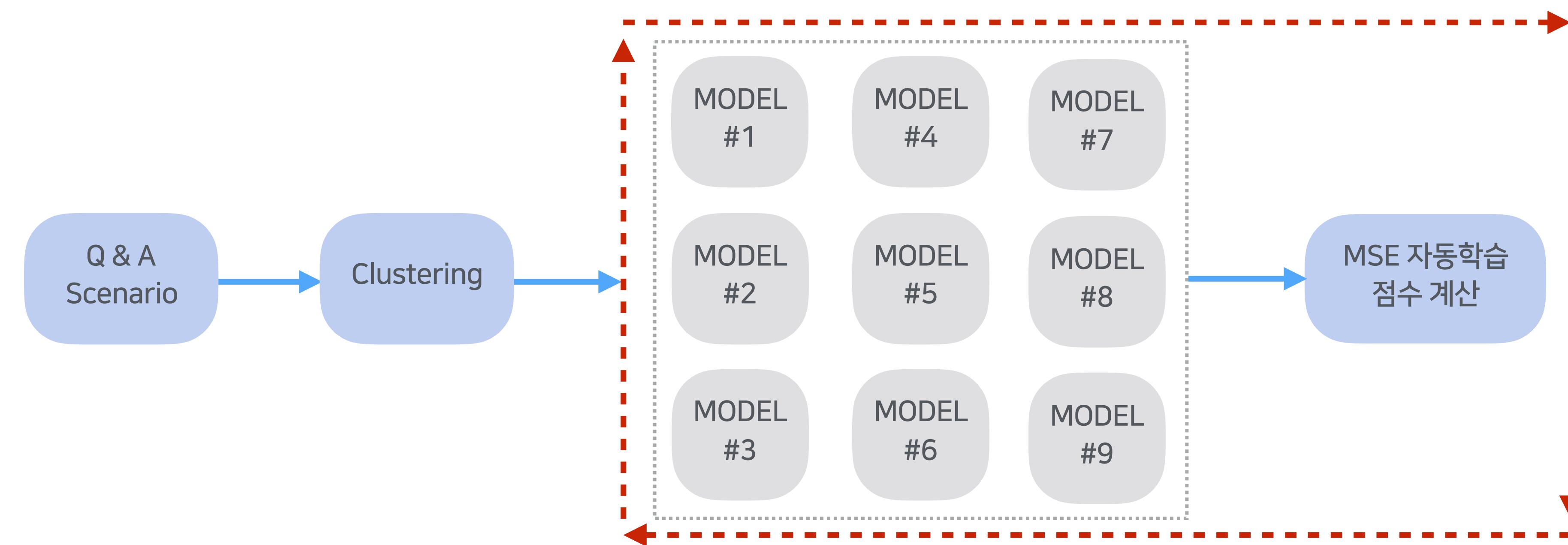
1. Dunn Index, Silhouette Score, MSE 등으로 모델 스스로 만족하는 시나리오 생성
2. 매번 새로운 Clustering을 할 때마다 Validation Set 제작
3. 자동학습 점수가 최대가 되는 Cluster k 선정



4.4 Hyper-Parameter 자동학습 방법

More Efficient, More Optimized

1. 그룹내 객체들이 정확하게 분류되었다는 가정
2. 말뭉치 또는 도메인에 따라 적절한 hyper-parameter 조합 상이
3. Cosine-Similarity 또는 Mean-Square Error로 평가 진행



4.5 코사인 유사도와 평균 제곱근 오차 (1/4)

DEVIEW
2018

간단하지만, 완벽했다

1. 자동평가 점수는 모델이 쿼리를 주어진 벡터 공간에 제대로 표현했는지 나타냄
2. 평균 제곱근 오차: 모델 학습을 위한 시나리오 & 모델의 학습 하이퍼 파라미터
3. 코사인 유사도: 모델이 학습 중 쿼리를 배치할 공간 (Embedding hidden sizes)

4.5 그렇다면 왜? (2/4)

코사인 유사도와 평균 제곱근 오차

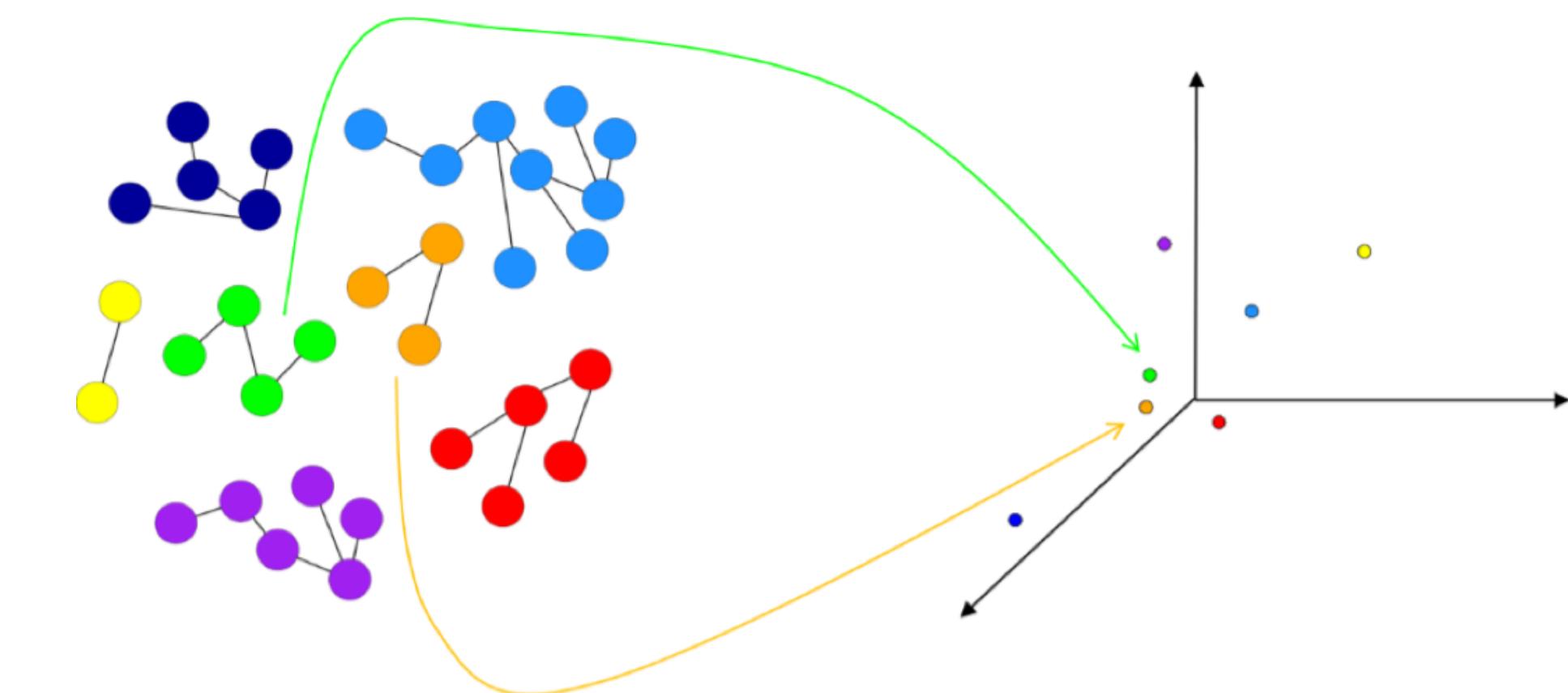
1. 코사인 유사도
 - A. 길이를 정규화해 비교하는 것과 유사
 - B. 벡터의 크기보다는 벡터 간 거리 측정하기 위한 메트릭
2. 평균 제곱근 오차
 - A. 시나리오에 얼마나 벗어났는지 regression 문제로 접근
 - B. Categorical value인 경우 Multinomial Distribution은 Maximum Entropy Distribution과 같음
 - C. 이를 Maximum likelihood estimation한 결과가 regression의 경우 MSE 최소화

5. 클러스터링, 그리고 파라미터 튜닝

5.1 What is Clustering?

“모델의, 모델에 의한, 모델을 위한” 클러스터링으로 기본 말뭉치를 챗봇 시나리오로 변환

1. 데이터에 대한 정보(label)없이 주어진 데이터를 가장 잘 설명하는 클러스터를 찾는 문제
2. EM Algorithm을 확장 및 응용하여 문제 해결
3. 갖고있는 정보를 최대한 그러나 효율적으로 활용하여 클러스터링
 - A. Feature extraction, if feature highly correlate, dimensionality reduction
 - B. PCA 또는 LDA(Linear Discriminant Analysis)



5.2 Gaussian Mixture Models (1/4)

DEVIEW
2018

확률과 확률분포를 기반으로 세우고 그것을 세우는 알고리즘

1. 데이터가 Gaussian 분포의 Mixture로 구성
2. k개의 Gaussian으로 전체 데이터를 표현
3. 데이터를 가장 잘 설명하는 k개의 평균과 공분산을 찾는 알고리즘
4. 클러스터 개수 k를 찾아주는 장점이 있으나, 문장의 embedding vector가 완벽하지 않다면 k를 제대로 형성하지 못하는 단점

5.2 DBScan Algorithm (2/4)

밀도 방식의 클러스터링 (Density-based spatial clustering of applications with noise)

1. 데이터가 세밀하게 몰려 있는 부분을 찾아서 클러스터링
2. 점 p로부터 거리 ϵ (epsilon)내에 점이 m개 있으면 하나의 군집으로 인식
3. 거리 ϵ 와 그 거리 안에 있는 점 m개를 parameter로 사용하는 알고리즘
4. 데이터 수가 적으면 제대로 작동하지만, 대부분의 경우 데이터 크기와 vector 공간 표현 제약 등 너무 sensitive하여 사용하기 어려움

5.2 Locality Sensitive Hashing (3/4)

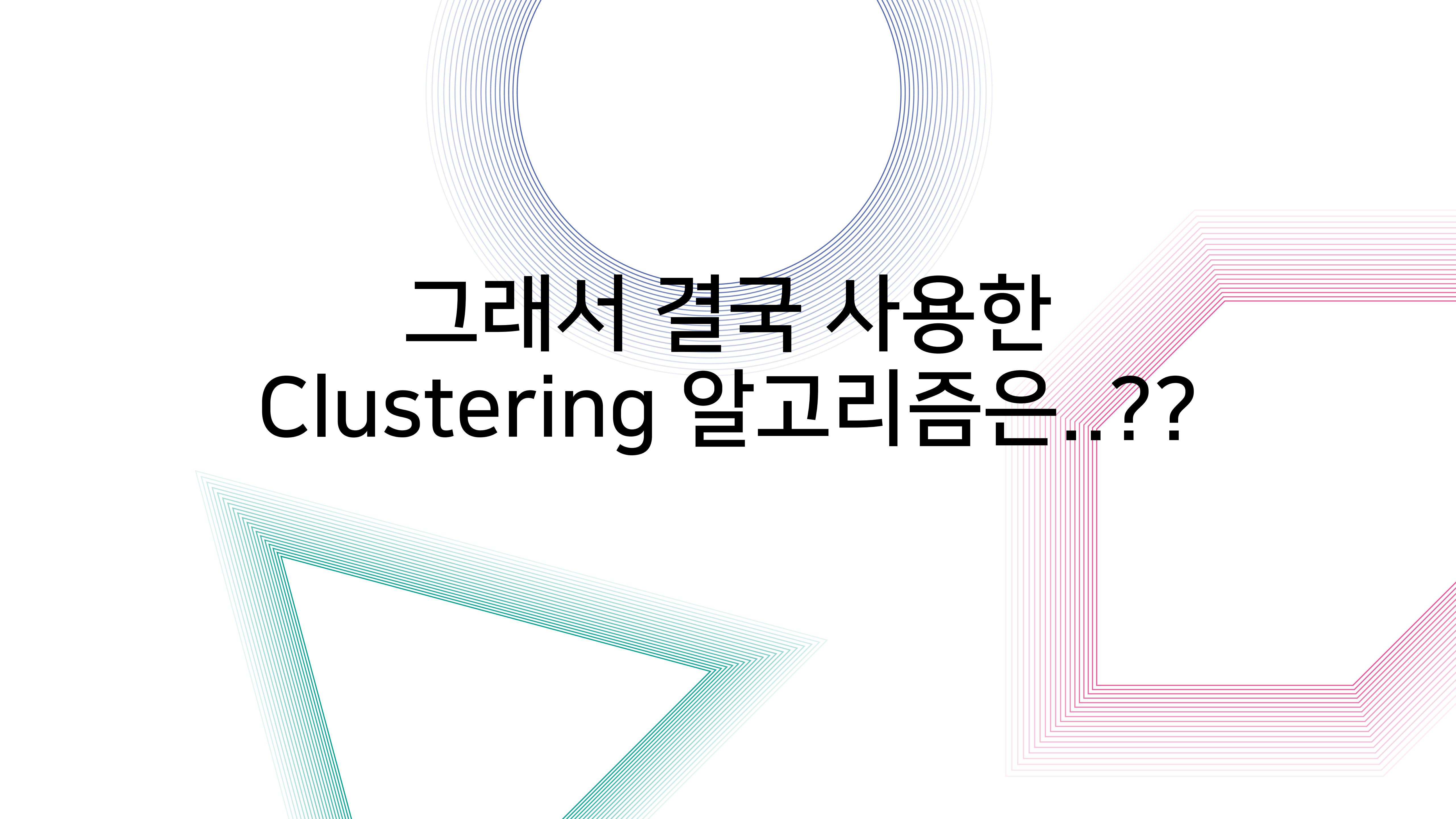
고차원 데이터의 차원 확률을 기반으로 차원 축소하여 유사도 계산 알고리즘

1. 클러스터링 하고자 하는 모든 쿼리에 대한 1:1 코사인 유사도를 구함
2. 정해둔 코사인 유사도 threshold이상인 쿼리끼리 묶어서 하나의 집합을 만듬
3. 각 쿼리의 pair끼리 코사인 유사도를 구하는 LSH 또는 ANN은 속도가 빠름
4. 단, A와 B, 그리고 B와 C가 코사인 유사도가 높다고 할 때, A와 C가 높지 않을 수 있으므로, List로 바꾸기 쉽지 않음

5.2 K-means Clustering (4/4)

최적화 함수를 거리 기반으로 세우고 그것을 푸는 알고리즘

1. 하나의 중심 centroid를 갖고, 각 개체는 가장 가까운 중심에 할당
2. 같은 중심을 이루는 개체들이 모여 하나의 군집을 형성
3. 답변보다 질의가 많다는 점에서 고유한 답변 개수를 초기 k로 설정



그래서 결국 사용한
Clustering 알고리즘은 ..??

5.3 역시 K-means Clustering!! (1/3)

DEVIEW
2018

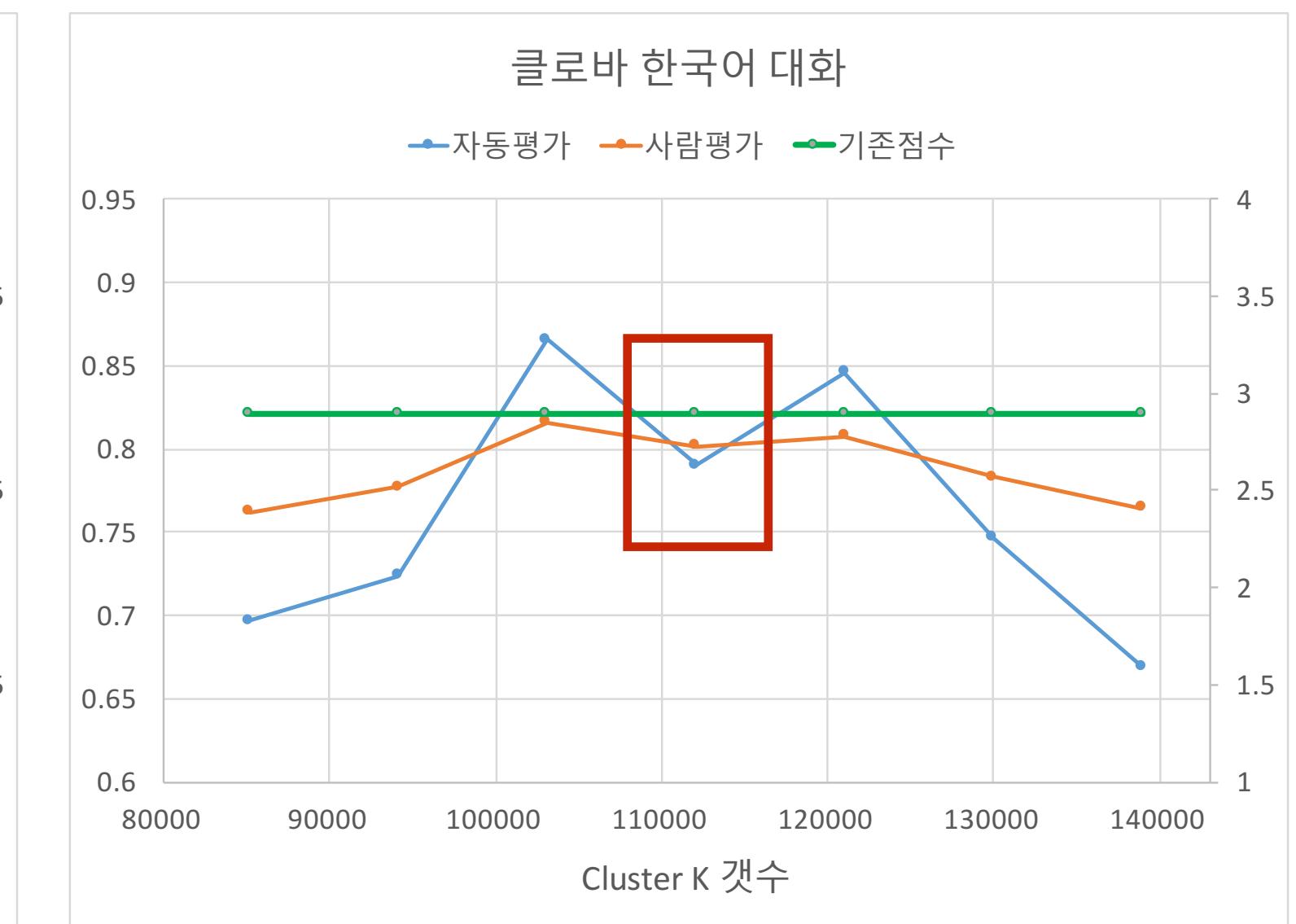
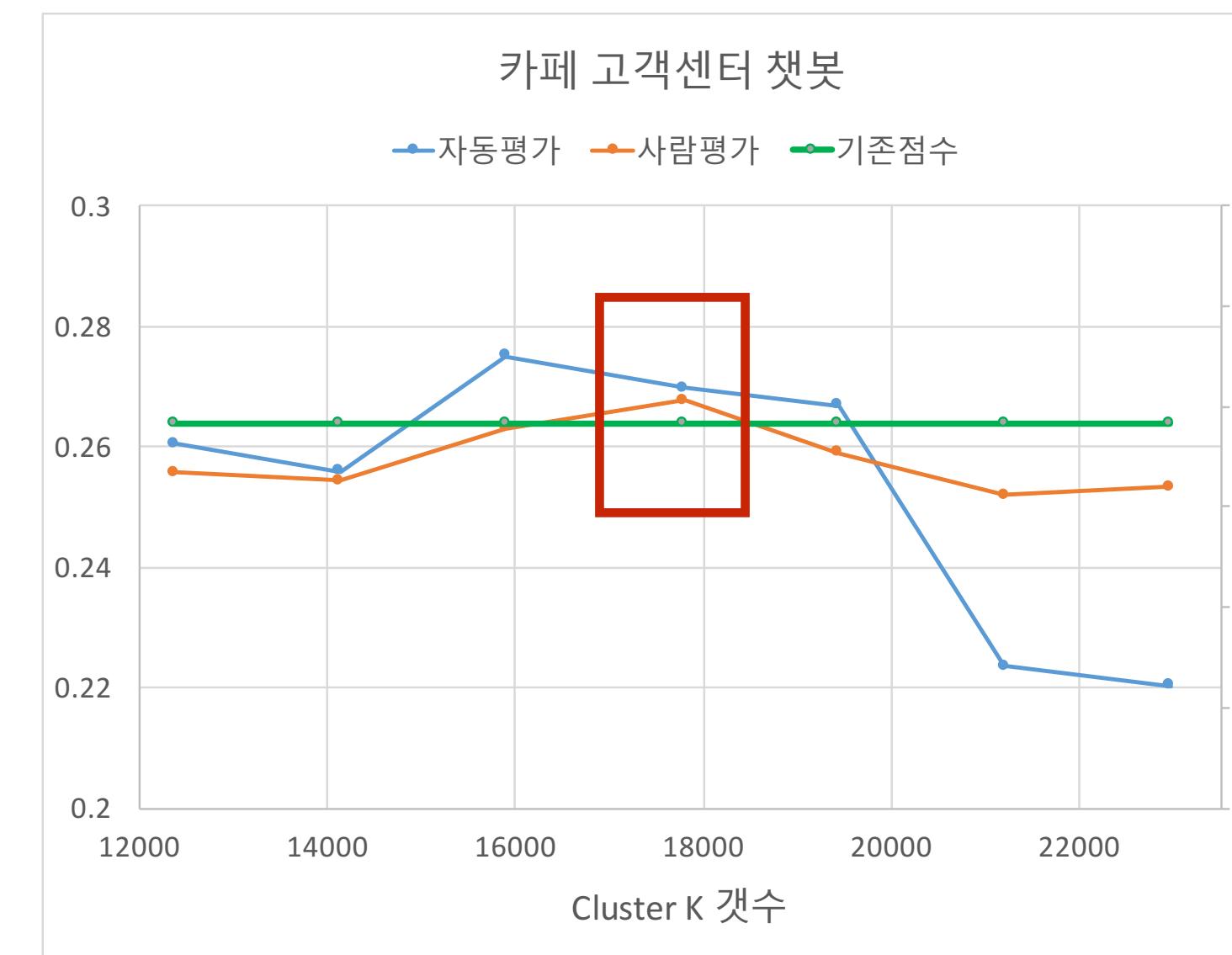
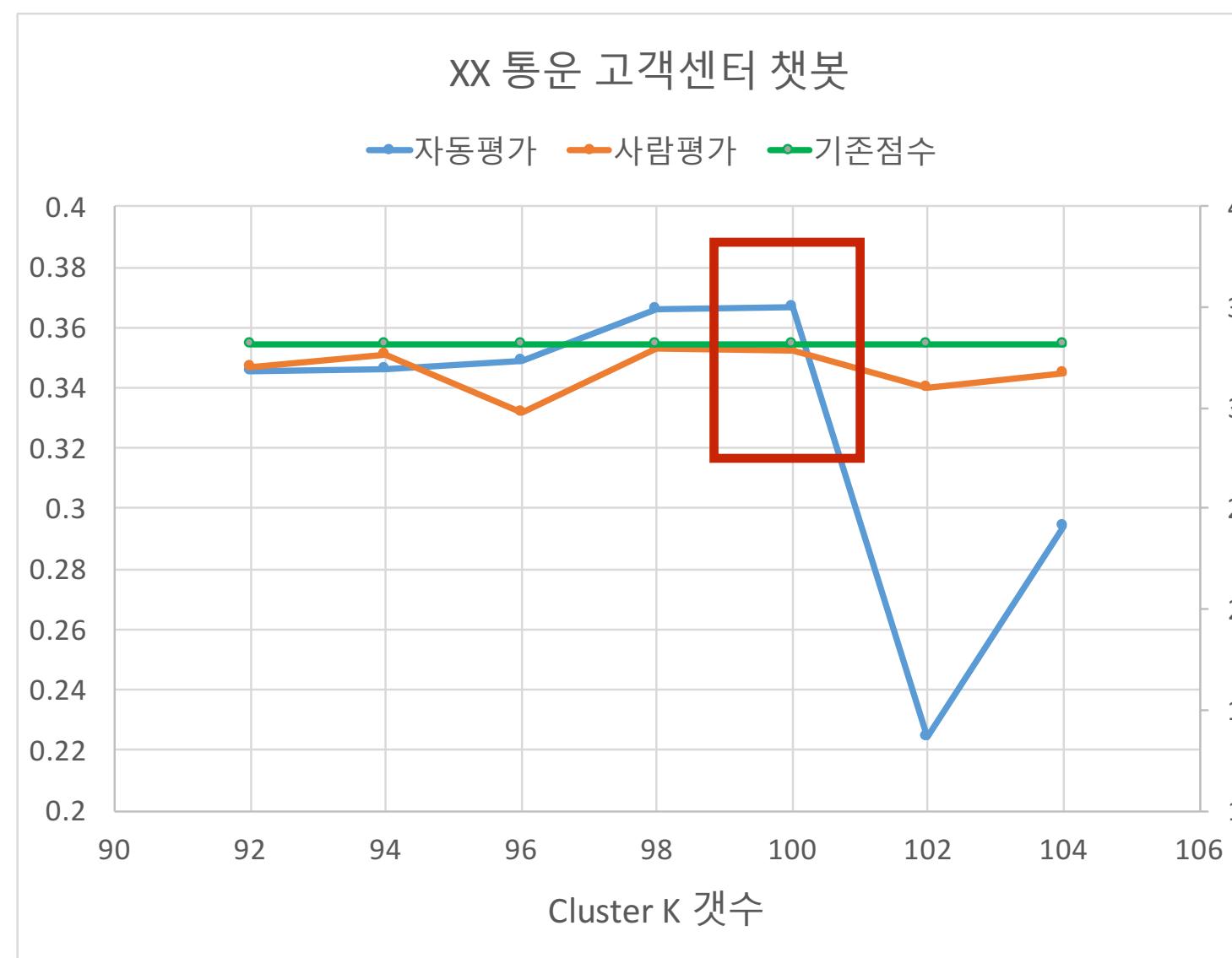
같은 문장임에도 속한 말뭉치마다 다르게 생성되는 embedding vector의 특징 반영

Data Input	▶	Data Output			
우유 맛나게 먹었냐		우유 맛나게 먹었냐	우유 맛나게 먹었니	우유 맛나게 먹었습니다	
우유 맛있는 거냐		우유 맛있는 거냐	우유 맛있는 거니		
우유 맛은 어때		우유 맛은 어때	우유 맛은 어때요	우유 맛은 어떠냐	우유 맛은 어떠니
우유 맛있네		우유 맛있네	우유 맛있다	우유 맛있습니다	우유맛있네
우유 맛나게 먹었니					우유맛있다
우유 맛있는 거니					우유맛있습니다
우유 맛은 어때요					우유맛있습니다
우유 맛있다					우유맛있습니다
우유 맛나게 먹었습니다					우유맛있습니다
우유 맛은 어떠냐					우유맛있습니다
우유 맛있습니다					우유맛있습니다
우유 맛은 어떠니					우유맛있습니다
우유맛있네					우유맛있습니다
우유 맛은 어떻다고					우유맛있습니다
우유맛있다					우유맛있습니다
우유 맛은 어떻다구					우유맛있습니다
우유맛있습니다					우유맛있습니다

5.3 역시 K-means Clustering!! (2/3)

DEVIEW
2018

그리하여,
많은 사람이 한땀 한땀 분류하고 시나리오 관리한 모델 성능과 비슷한 성능의 시나리오를 만들었다



기준 그룹 수: 100, 기준 쿼리 수: 661
총 clustering 횟수: 9번

기준 그룹 수: 17,668, 기준 쿼리 수: 61,182
총 clustering 횟수: 23번

기준 그룹 수: 112,067, 기준 쿼리 수: 624,751
총 clustering 횟수: 72번

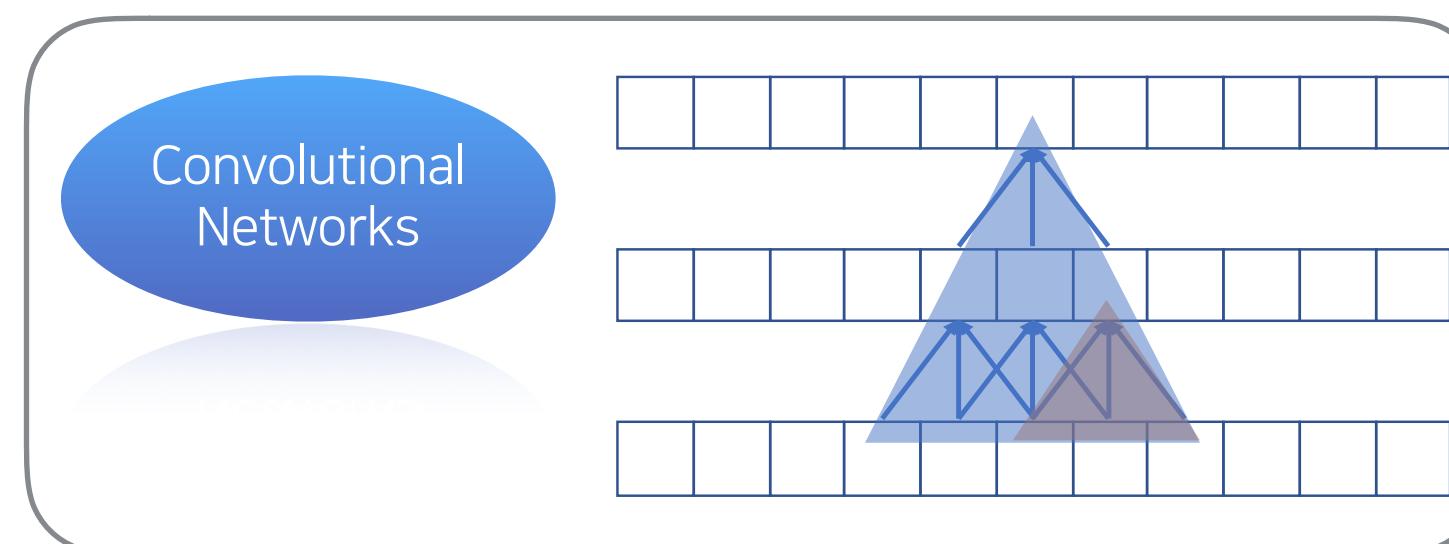
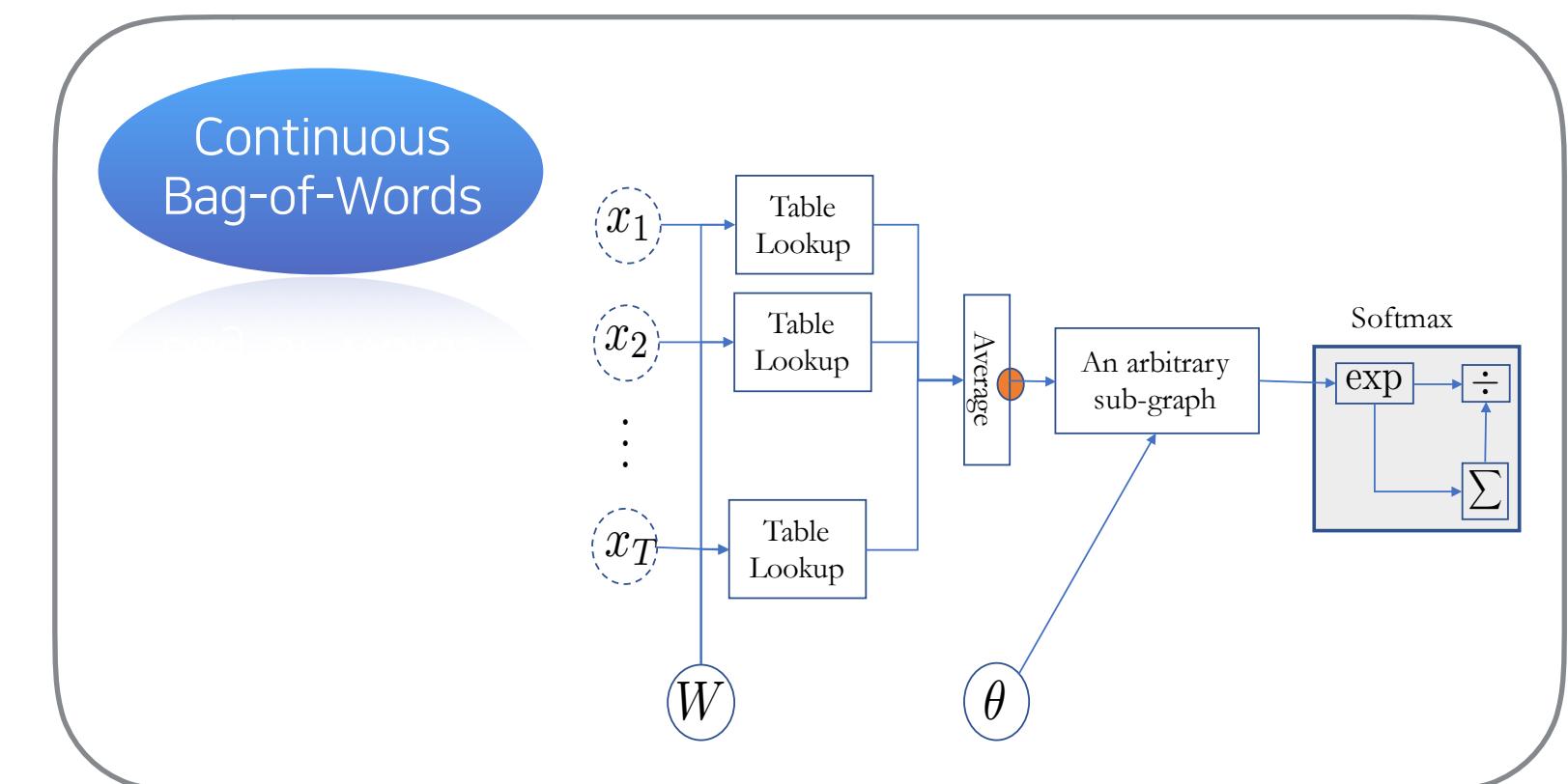
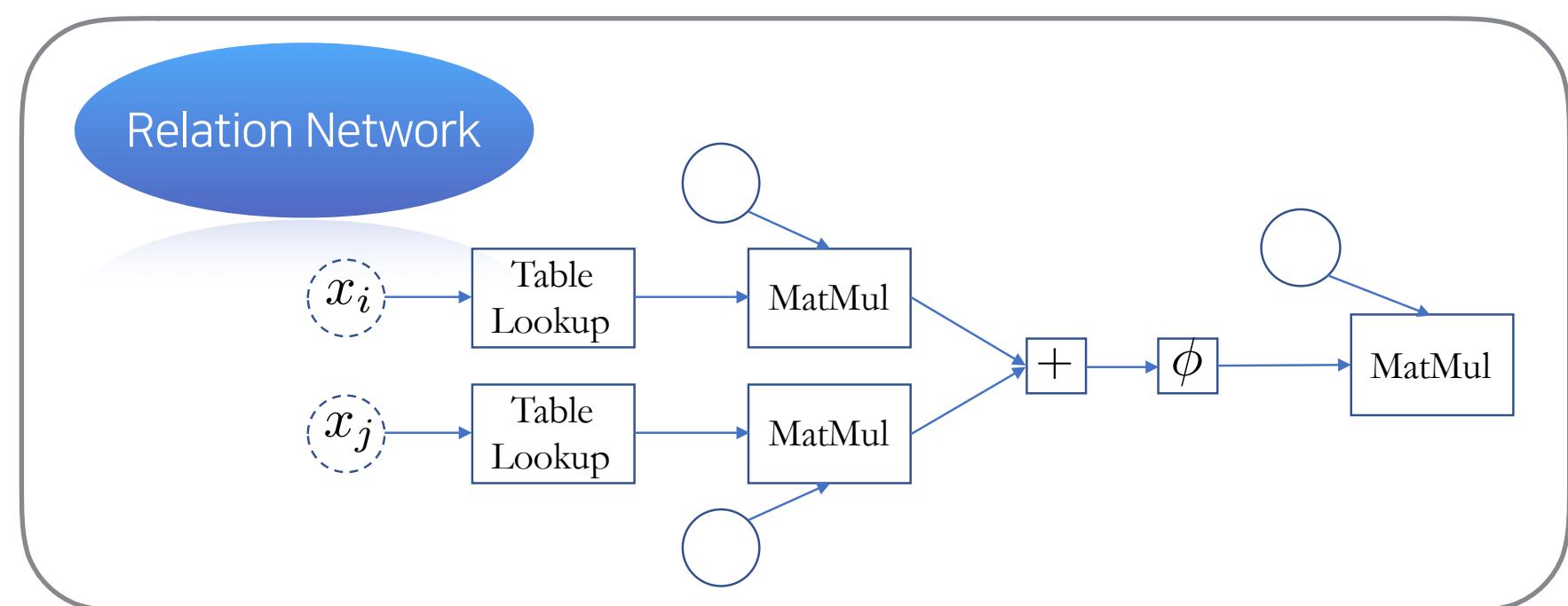
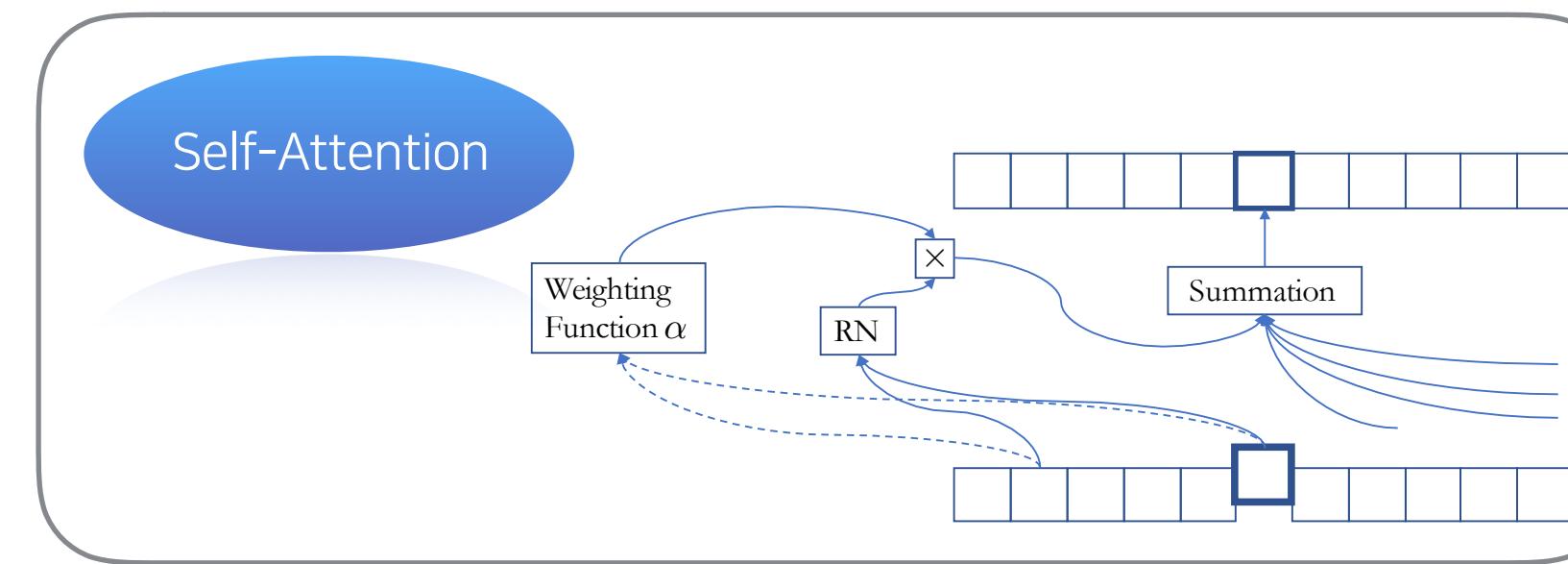
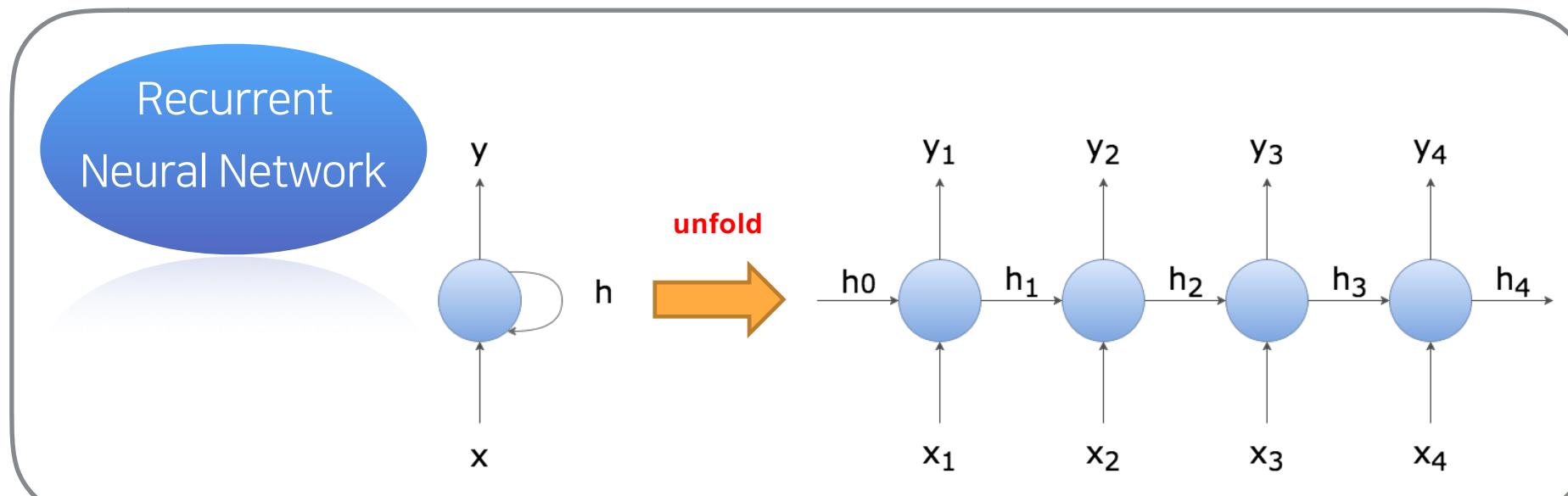
5.3 역시 K-means Clustering!! (3/3)

Lexical Variance in Queries

1. 서로 다른 렉시컬 쿼리를 하나의 의도라고 묶어 주는 것이 중요
2. 하나의 의도에 다양한 variation이 학습되면 챗봇 커버리지 증가
3. 쿼리를 원형복원(Lemmatization)하여 의도별로 분류하는 것이 중요

이번에는
최적의 하이퍼 파라미터 조합을
찾아볼까요??

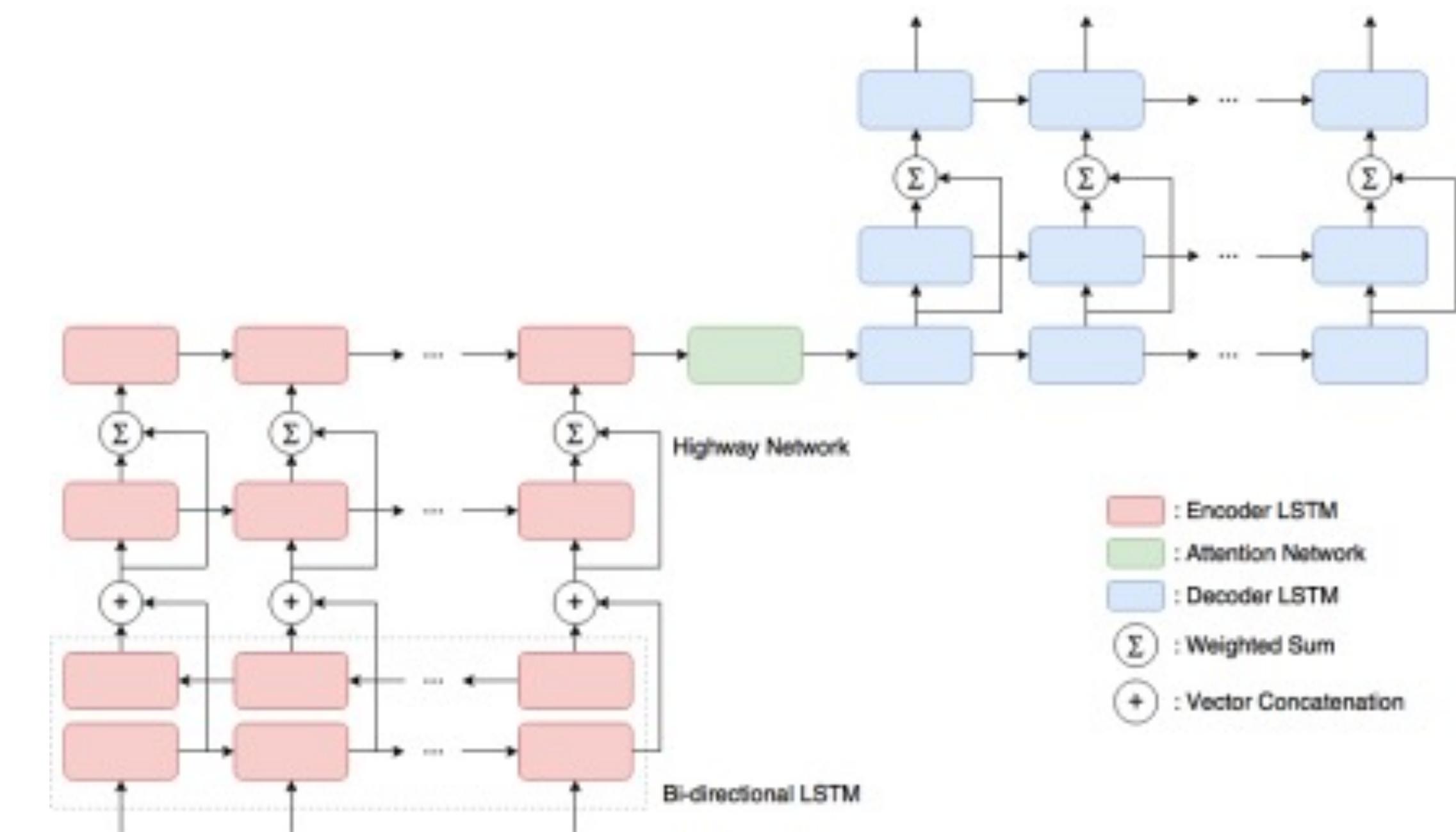
5.4 자연어처리를 위한 딥러닝 모델



5.5 하지만 우리의 선택은? (1/2)

Seq2Seq-based Retrieval Model

→ Google의 GMNT에서 아이디어를 착
안하여 한국어 및 대화형 데이터에 최적인
대화 모델 설계



5.5 Simple but Satisfied Model (2/2)

Sequence-To-Sequence Model Hyper-parameters

1. 기본 모델도 최적의 하이퍼 파라미터 조합으로 충분히 성능 향상 기대
2. 그러나, 시간과 학습에 필요한 GPU 등 상당한 양의 리소스가 요구됨

Type	Default Value	Type	Default Value
Max Epoch	100000	Dropout Input Keep Probability	0.8
Batch Size	1024	Dropout Intra Keep Probability	0.5
Unknown Count	3	Dropout Output Keep Probability	0.8
Target Perplexity	6.5	Dropout State Keep Probability	0.75
Threshold	0.963	Buckets	[8, 13]
Unk Ratio	0.333	Hidden Sizes	[64, 64, 32] & [32, 64, 128]
Learning Rate	0.0005	Bottleneck	FALSE
Unknown Percent	0.33	Bidirectional	TRUE
Max Gradient Norm	5	Min Vocab Count	2

5.6 Model Hyper-parameters 예시 (2/2)

DEVIEW
2018

말뭉치 사이즈별 Model Hyper-Parameters 예시

	클로바 FAQ 챗봇	XXX 클라우드 챗봇	클로바 한국어대화
쿼리 수	662,917건	45,941건	670,374 건
시나리오 수	52건	25,187건	112,067건
hiddenSizes Vector	24	28	52
Learning Rate	0.001	0.001	0.00025
minVocab Count	0	2	4

Type	Value	MSE	Cosine Similarity
dropout_input_keep_prob	0.56 2018-07-19T08-19-32.184Z	0.689266826	0.745406392
dropout_input_keep_prob	0.64 2018-07-19T08-21-00.426Z	0.720242796	0.779333711
dropout_input_keep_prob	0.72 2018-07-19T08-22-30.981Z	0.708121628	0.770150177
dropout_input_keep_prob	0.8 2018-07-19T08-24-00.866Z	0.704501768	0.757599383
dropout_input_keep_prob	0.88 2018-07-19T08-25-36.067Z	0.743370536	0.807477333
dropout_input_keep_prob	0.96 2018-07-19T08-27-05.512Z	0.721603122	0.789145925
dropout_input_keep_prob	1 2018-07-19T08-28-35.765Z	0.741947532	0.813354366
dropout_intra_keep_prob	0.35 2018-07-19T08-30-05.68Z	0.728895982	0.771065331
dropout_intra_keep_prob	0.4 2018-07-19T08-31-36.105Z	0.735780359	0.772712363
dropout_intra_keep_prob	0.45 2018-07-19T08-33-05.809Z	0.725963128	0.788949325
dropout_intra_keep_prob	0.5 2018-07-19T08-34-35.786Z	0.707691344	0.765514083
dropout_intra_keep_prob	0.55 2018-07-19T08-36-05.156Z	0.695002165	0.755692942
dropout_intra_keep_prob	0.6 2018-07-19T08-37-34.999Z	0.737243227	0.786318732
dropout_intra_keep_prob	0.65 2018-07-19T08-39-05.225Z	0.732347208	0.786633738
dropout_state_keep_prob	0.525 2018-07-19T08-40-35.369Z	0.631561321	0.702254171
dropout_state_keep_prob	0.6 2018-07-19T08-42-05.32Z	0.719987182	0.78432071
dropout_state_keep_prob	0.675 2018-07-19T08-43-35.159Z	0.723206358	0.767377793
dropout_state_keep_prob	0.75 2018-07-19T08-45-06.015Z	0.723791699	0.77116016
dropout_state_keep_prob	0.825 2018-07-19T08-46-35.477Z	0.727184012	0.79018914
dropout_state_keep_prob	0.9 2018-07-19T08-48-20.073Z	0.69946665	0.77366472
dropout_state_keep_prob	0.975 2018-07-19T08-49-52.846Z	0.656554135	0.723204238
dropout_output_keep_prob	0.56 2018-07-19T08-51-24.157Z	0.70571072	0.765265171
dropout_output_keep_prob	0.64 2018-07-19T08-52-53.697Z	0.747159588	0.783743363
dropout_output_keep_prob	0.72 2018-07-19T08-54-20.737Z	0.732814961	0.799940322
dropout_output_keep_prob	0.8 2018-07-19T08-55-49.116Z	0.686932261	0.745503589
dropout_output_keep_prob	0.88 2018-07-19T08-57-16.787Z	0.77055466	0.814460178
dropout_output_keep_prob	0.96 2018-07-19T08-58-49.349Z	0.697745601	0.765938774
dropout_output_keep_prob	1 2018-07-19T09-00-17.918Z	0.708872834	0.765719429



Type	MSE	Cosine Similarity
best	2018-07-23T04-57-48.199Z	0.146264675
original	2018-07-20T06-22-48.945Z	0.411612663

5.6 그리고 실제 결과 (2/2)

차원 수	20	24	28	32	36	44	48	52	56	64	68	72	76	80	88	96	104	112	128
코사인 유사도 점수	0.00241	0.00373	0.00236	0.00977	0.00598	0.00325	0.01231	0.02275	0.00433	0.00562	0.00681	0.00646	0.00762	0.00821	0.00945	0.01089	0.00809	0.01020	0.01464

적절한 Vector Space 찾기

1. 기존 128차원보다 자동학습으로 찾은 52차원이 **약 6%** 성능 향상
2. 전체 말뭉치: 약 68만 여건
3. 전체 시나리오: 약 11만 여건

평가	원본 128차원		수정 52차원	
	건수	%	건수	%
1점	628건	22.8%	492건	17.9%
2점	317건	11.5%	274건	10.0%
3점	531건	19.3%	541건	19.7%
4점	1,275건	46.3%	1,445건	52.5%
총합계	2,751건	100%	2,752건	100%
평가점수	2.89점		3.07점	

6. 최종 모델 성능 및 상용 서비스 현황

6.1 USER에게 BUILD란…? (1/4)

DEVIEW
2018

STEP 1. 대화 시나리오 생성

The screenshot shows the DEVIEW 2018 Chatbot interface. On the left, there's a sidebar with sections like '챗봇 설정' (Chatbot Settings), '빌드 관리' (Build Management), '대화 생성' (Conversation Generation), '대화 시나리오' (Conversation Scenario) which is currently selected and has a blue background, '대화 유형' (Conversation Type), '엔티티' (Entity), '답변 콤포넌트' (Response Component), '액션 메소드' (Action Method), '폼' (Form), and '태스크' (Task). The main area is titled '대화 시나리오 ②' and shows a list of conversation scenarios. The first scenario is '로그인이 안돼요' (Login is not available) with a '로그인' (Login) button. The second is '아이디가 기억 나지 않아요' (I don't remember my ID) with a '로그인' (Login) button. The third is '비밀번호가 기억 나지 않아요' (I can't remember my password) with a '로그인' (Login) button. The fourth is '안녕하세요' (Hello) with an '안녕' (Hello) button. Each scenario includes a question, a response, and a timestamp. There are buttons for '저장' (Save) and '수정' (Edit) next to each row. At the bottom, there are navigation arrows and a '10개씩 보기' (View 10 at a time) button.

ID	시나리오 이름	대화 유형	질문	답변	작성/변경일 (UTC+09:00)
1698358	로그인이 안돼요	로그인	로그인이 안돼요	로그인이 되지 않는 경우 가장 쉽고 빠르게 해결할 수 있는 방법은 먼저 아이디, 비밀번호 찾기를 하는 것입니다. ...	2018-07-13 17:05:34
1698357	아이디가 기억 나지 않아요	로그인	아이디를 모르겠어요	이용자의 개인정보 보호를 위해 아이디 소유자 정보와 일치하는지 본인 인증 절차가 필요합니다. 로그인 페이...	2018-07-13 17:05:34
1698356	비밀번호가 기억 나지 않아요	로그인	비번을 모르겠어요	이용자의 개인정보 보호를 위해 아이디 소유자 정보와 일치하는지 본인 인증 절차가 필요합니다. 로그인 페이...	2018-07-13 17:05:34
1698355	안녕하세요	안녕	안녕이란 무엇일까	이건 안녕에 대한 답변입니다	2018-07-13 16:13:35

6.1 USER에게 BUILD란…? (2/4)

DEVIEW
2018

STEP 2. USER가 채팅할 것 같은 질문과 챗봇이 보여줄 답변 등록

질문 등록
다양한 사용자 질문을 등록하세요. 입력된 질문 목록에서 드래그로 엔티티등록이 가능합니다.

질문 메시지 작성 *	다양한 사용자 표현을 입력하고 추가를 눌러주세요	+ 추가
질문4	deview 행사 장소는 어디인가요?	수정 삭제
질문3	데뷰는 언제하나요?	수정 삭제
질문2	DEVIEW 2018 시간과 장소는 어떻게 되나요?	수정 삭제
질문1	DEVIEW 2018 행사일은 어떻게 되나요?	수정 삭제

답변 등록 랜덤 ②
사용자에게 전달할 답변을 등록할 수 있습니다.(• 필수 입력 사항입니다.)

답변 메시지 작성 *	답변을 입력하세요. \$ 입력하면 액션 메소드 리스트가 보여지며 선택시 액션메소드의 형태로 답변에 추가됩니다. # 입력하면 품 리스트가 보여지며 선택시 품의 형태로 답변에 추가됩니다.	컴포넌트 추가
서비스 ON/OFF설정	<input checked="" type="checkbox"/>	
URL 주소	예시) http://ncloud.com/compute/information	
노출문구	사용자에게 URL주소대신 노출할 문구를 입력하세요.	
+ 답변 추가		
답변 1	답변 메시지 수정 삭제	답변 메시지 수정 삭제
서비스 여부	<input checked="" type="checkbox"/>	네, 발표자료는 세션 시작 바로 전 DEVIEW 홈페이지에 게재될 예정으로, 발표내용을 들으시면서 자료를 보실 수 있으며 발표영상은 행사 이후 DEVIEW 홈페이지에 업로드될 예정입니다.

저장

질문 등록
다양한 사용자 질문을 등록하세요. 입력된 질문 목록에서 드래그로 엔티티등록이 가능합니다.

질문 메시지 작성 *	다양한 사용자 표현을 입력하고 추가를 눌러주세요	+ 추가
질문4	발표자료는 홈페이지에 올라오나요?	수정 삭제
질문3	세션을 놓쳤는데 어디서 다시 봐요?	수정 삭제
질문2	세션별 발표자료와 영상은 다시 볼 수 있나요?	수정 삭제
질문1	각 세션별 발표자료와 영상은 어디서 보나요?	수정 삭제

답변 등록 랜덤 ②
사용자에게 전달할 답변을 등록할 수 있습니다.(• 필수 입력 사항입니다.)

답변 메시지 작성 *	답변을 입력하세요. \$ 입력하면 액션 메소드 리스트가 보여지며 선택시 액션메소드의 형태로 답변에 추가됩니다. # 입력하면 품 리스트가 보여지며 선택시 품의 형태로 답변에 추가됩니다.	컴포넌트 추가
서비스 ON/OFF설정	<input checked="" type="checkbox"/>	
URL 주소	예시) http://ncloud.com/compute/information	
노출문구	사용자에게 URL주소대신 노출할 문구를 입력하세요.	
+ 답변 추가		
답변 1	답변 메시지 수정 삭제	답변 메시지 수정 삭제
서비스 여부	<input checked="" type="checkbox"/>	네, 발표자료는 세션 시작 바로 전 DEVIEW 홈페이지에 게재될 예정으로, 발표내용을 들으시면서 자료를 보실 수 있으며 발표영상은 행사 이후 DEVIEW 홈페이지에 업로드될 예정입니다.

저장

6.1 USER에게 BUILD란...? (3/4)

DEVIEW
2018

STEP 3-1. 대화 모델 빌드 요청/관리 및 배포

The screenshot shows the DEVIEW 2018 platform interface. At the top, there is a navigation bar with icons for messaging, service deployment, and user profile. Below the navigation bar, the main area is divided into two sections:

- Left Section (Chatbot Management):** This section contains a sidebar with various options like '대화 생성', '대화 시나리오', '대화 유형', '엔티티', '답변 콤보넌트', '액션 메소드', '폼', '태스크', '테스트 및 기록', and '테스트'. The '대화 시나리오' section is currently selected, indicated by a blue plus sign icon.
- Right Section (Build Request):** This section has tabs for '챗봇 설정' (Chatbot Settings), '도메인정보' (Domain Information), '메신저 연동' (Messenger Integration), and '금칙어 관리' (Profanity Management). The '챗봇 설정' tab is active. A modal window titled '대화모델 빌드' (Conversation Model Build) is open, asking '대화 모델 빌드를 시작하시겠습니까?' (Do you want to start the conversation model build?). It contains two bullet points:
 - 시나리오 개수가 많을 수록 더 많은 빌드 시간이 소요됩니다. (More scenarios mean longer build times.)
 - 대화 모델 빌드가 시작되면 완료 시까지 중지할 수 없습니다. (The build cannot be stopped once it starts until completion.)
- Bottom Section (Build History):** This section lists build logs with columns for '종료 시간' (Completion Time), '현재 서비스 여부' (Current Service Status), and '동작' (Actions). Two specific rows are highlighted with red boxes:
 - Row 1: 종료 시간 - 2018-10-10 17:18:05, 현재 서비스 여부 - Service, 동작 - 베타 배포, 서비스 배포
 - Row 2: 종료 시간 - 2018-10-10 17:19:05, 현재 서비스 여부 - Beta, 동작 - 베타 배포, 서비스 배포

6.1 USER에게 BUILD란…? (4/4)

DEVIEW
2018

STEP 4. 학습완료된 빌드로 챗봇 테스트

The screenshot shows the DEVIEW 2018 Chatbot interface. The top navigation bar includes 'Chatbot' (with a globe icon), '대화 모델 빌드' (Conversation Model Build), '서비스 배포' (Service Deployment), and user profile icons.

The main area is titled '대화 기록' (Conversation Log) with a search bar. It displays a list of messages from four users labeled TEST 1 through TEST 4. The messages are:

- TEST 1: 아이디를 모르겠어 (UTC+09:00)
- TEST 2: 너 누구야? (UTC+09:00)
- TEST 3: 반갑습니다 (UTC+09:00)
- TEST 4: 누구니 (UTC+09:00)

A large callout bubble from TEST 1 contains a note: "A 이용자의 개인정보 보호를 위해 아이디 소유자 정보와 일치하는지 본인 인증 절차가 필요합니다. 로그인 페이지 아래 '아이디 찾기'에서 회원정보로 등록한 이메일 주소 또는 휴대전화번호로 본인 인증 후 아이디를 확인하실 수 있어요." (For user privacy protection, it is necessary to verify the user's identity by checking if the registered email address or mobile phone number matches the owner's information. You can do this by logging in and using the 'Forgot ID' feature.)

The left sidebar, titled 'test', contains sections for '대화 생성' (Conversation Generation) with '대화 시나리오' (+), '대화 유형', and '엔티티'; '답변 콤포넌트' (Response Component) with '액션 매소드', '폼', and '태스크'; and '테스트 및 기록' (Test and Log) with '테스트' (checked) and '대화 기록' (highlighted with a red box).

A red box highlights the '테스트' button in the sidebar and the '대화 기록' link in the log section. Another red box highlights the '테스트' button at the bottom right of the screen.

6.2 사용자는 모르지만... (1/2)

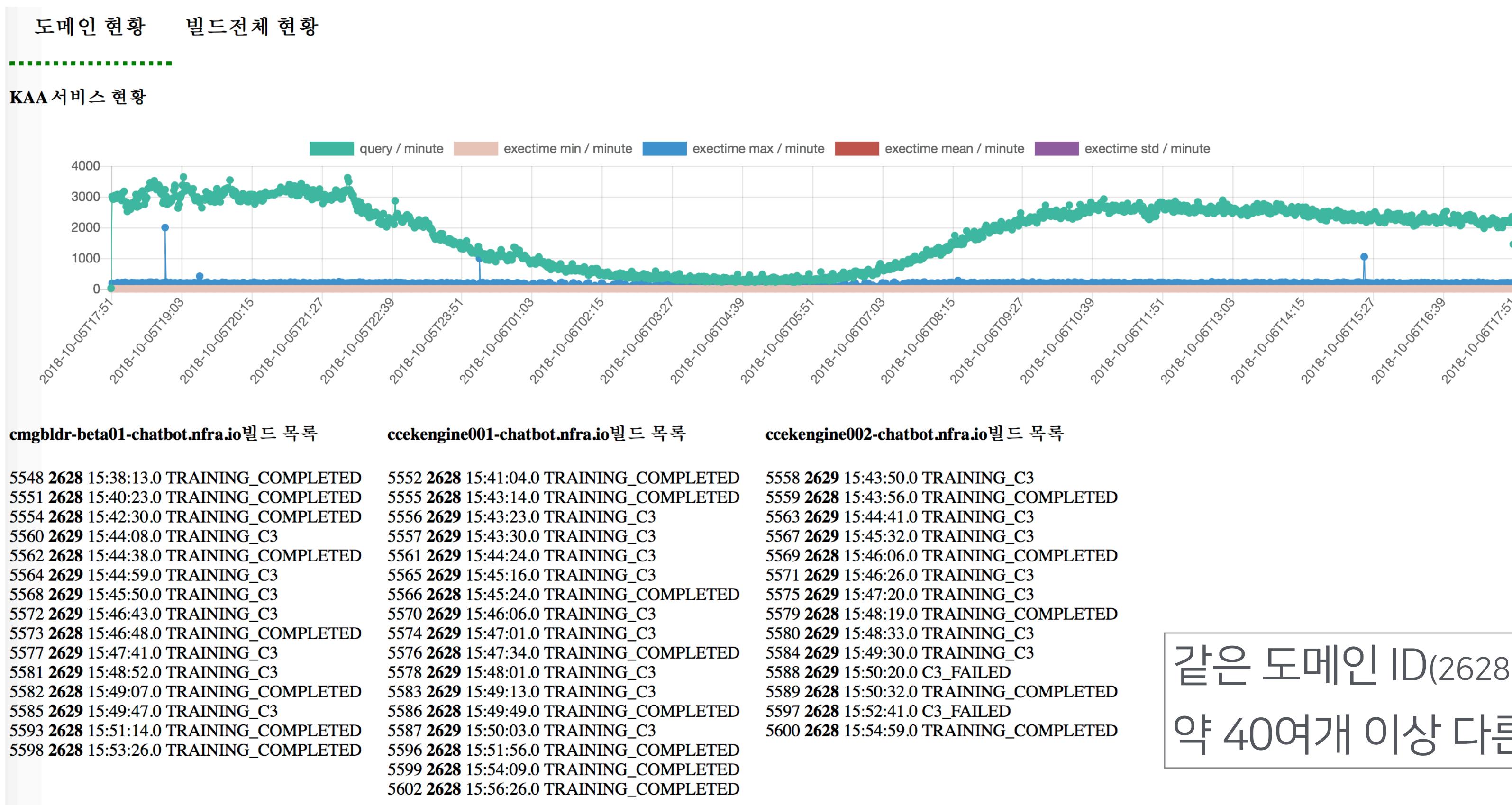
Build 요청 후 시간이 흐르면 Chatbot 등장

1. 일반적인 Chatbot Build를 진행
2. 하지만 1차 Build 완료와 동시에 AutoML로 back-ground update 실시
3. 모델이 스스로 만든 시나리오와 스스로 최적화한 하이퍼 파라미터로 최종 모델 제안

6.2 여전히 사용자는 모르고... (2/2)

DEVIEW
2018

더 나은 챗봇을 서비스하기 위한 추가 학습이 진행되는 과정



6.3 다양한 모델 적용가능?

Tensorflow로 구현된 모델이라면 얼마든지 가능

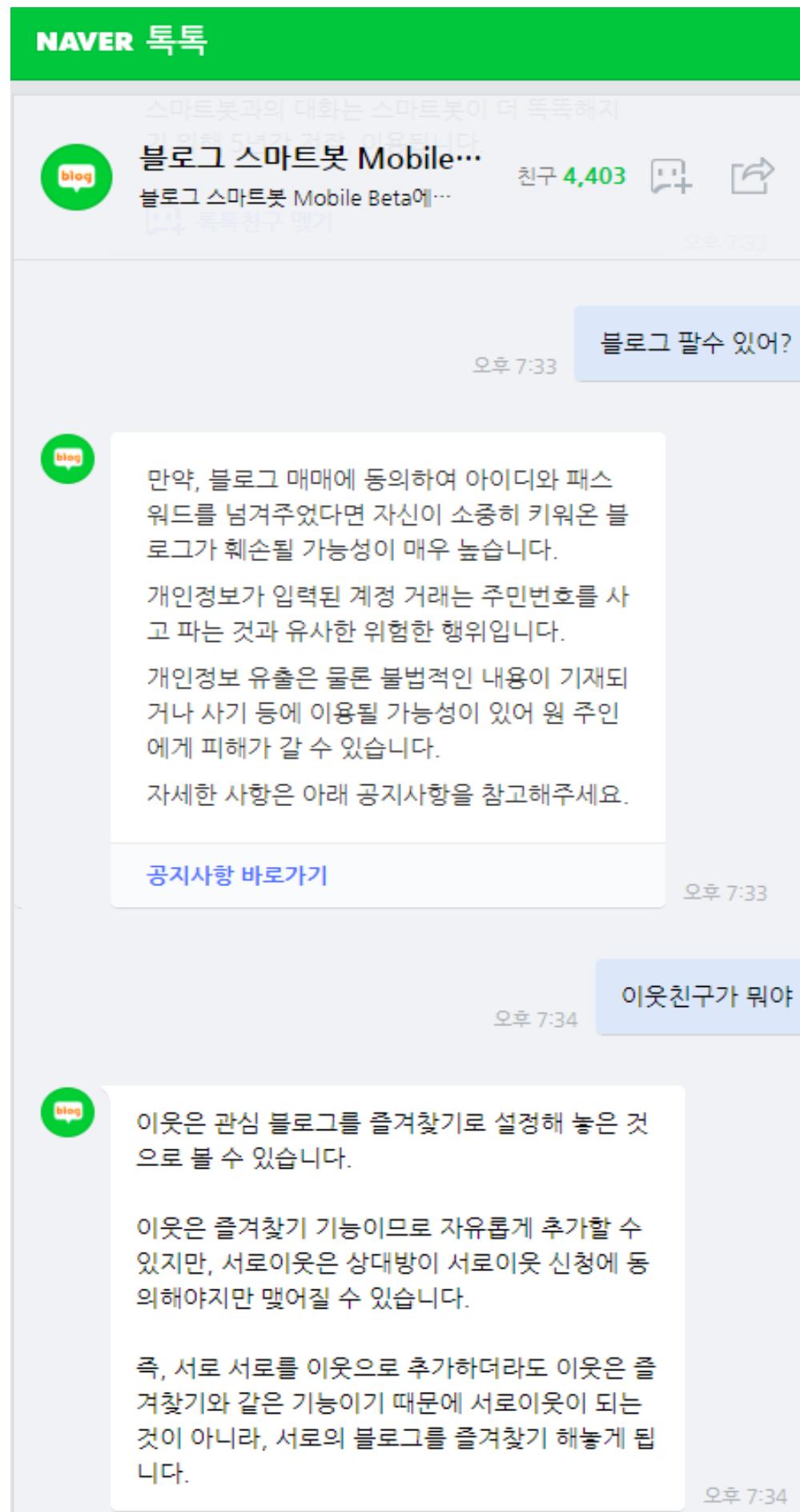
1. 모델 jar파일 업로드
2. 풍부한 컴퓨터 리소스 자원으로 Spark 분산 처리 및 NAVER c3dl GPU로 학습 진행
3. LSTM으로 구현된 예시 모델이 아닌 Attention 모델 등 다양한 모델 가능

6.4 실제 서비스 현황 (1/3)

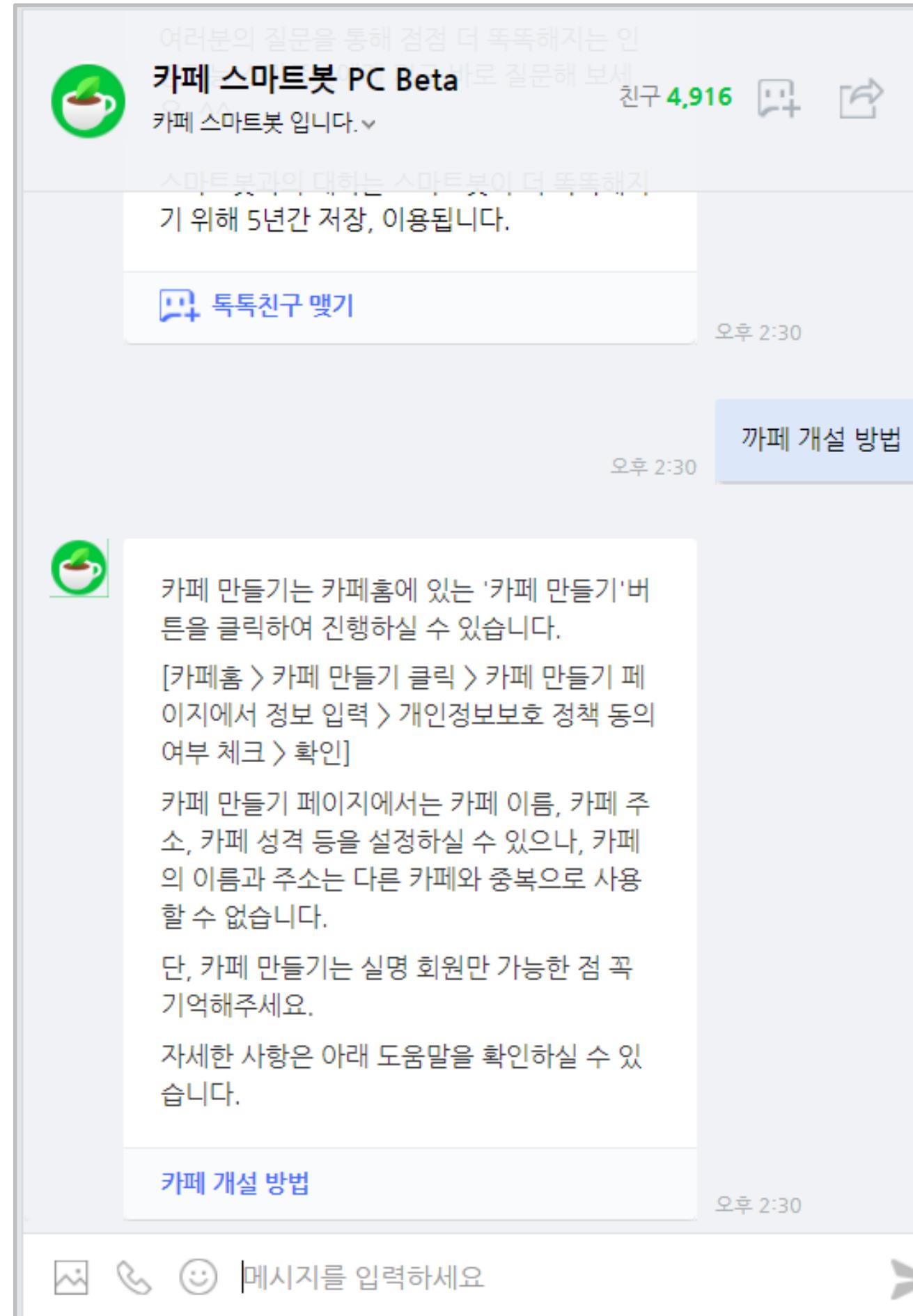
현재 실서비스하는 도메인 중 대다수 AutoML 지원

1. 한국어 대화 등 클로바 다수 도메인
2. 네이버 비즈니스 플랫폼 클라우드봇 (NAVER Cloud Platform)
 - A. 주요 클라우드 용어 및 상품명을 Entity Tagging 실시
 - B. 답변하지 못하는 데이터 지속적으로 재학습 진행
3. NAVER 블로그, 카페 등 고객센터 챗봇
 - A. 서비스 적용된 네이버 고객센터 챗봇 현재 9개
 - B. 블로그 도움말 페이지 방문 횟수 **평균 15% 감소**
 - C. Modoo 고객센터 문의건수 **약 25% 감소**
 - D. 현재 상용 Beta 서비스 중. 하루 기준 700~1000명, 2천~3천 발화 대응중

6.4 실제 서비스 현황 (2/3)



블로그 고객센터 예시 (PC)

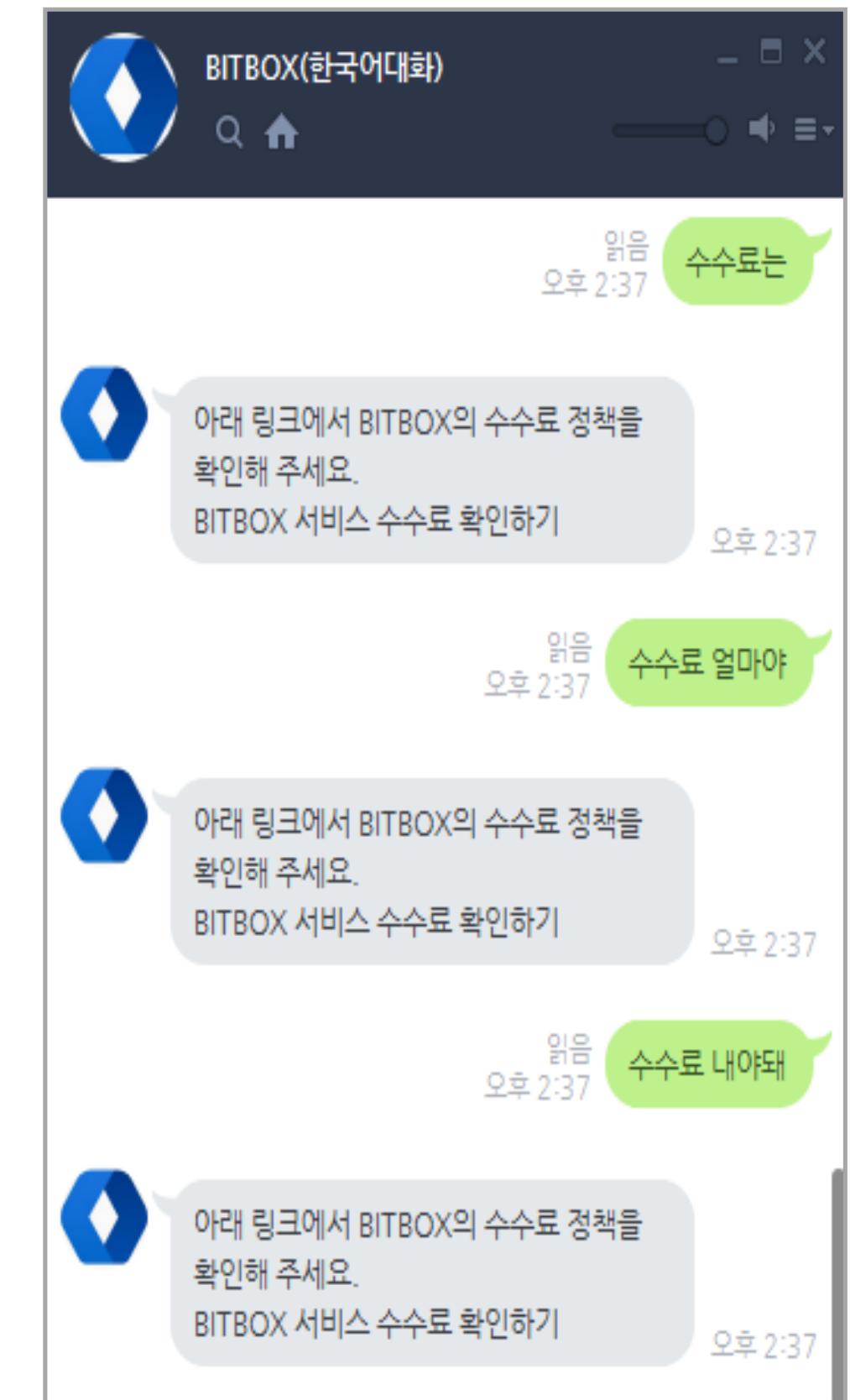
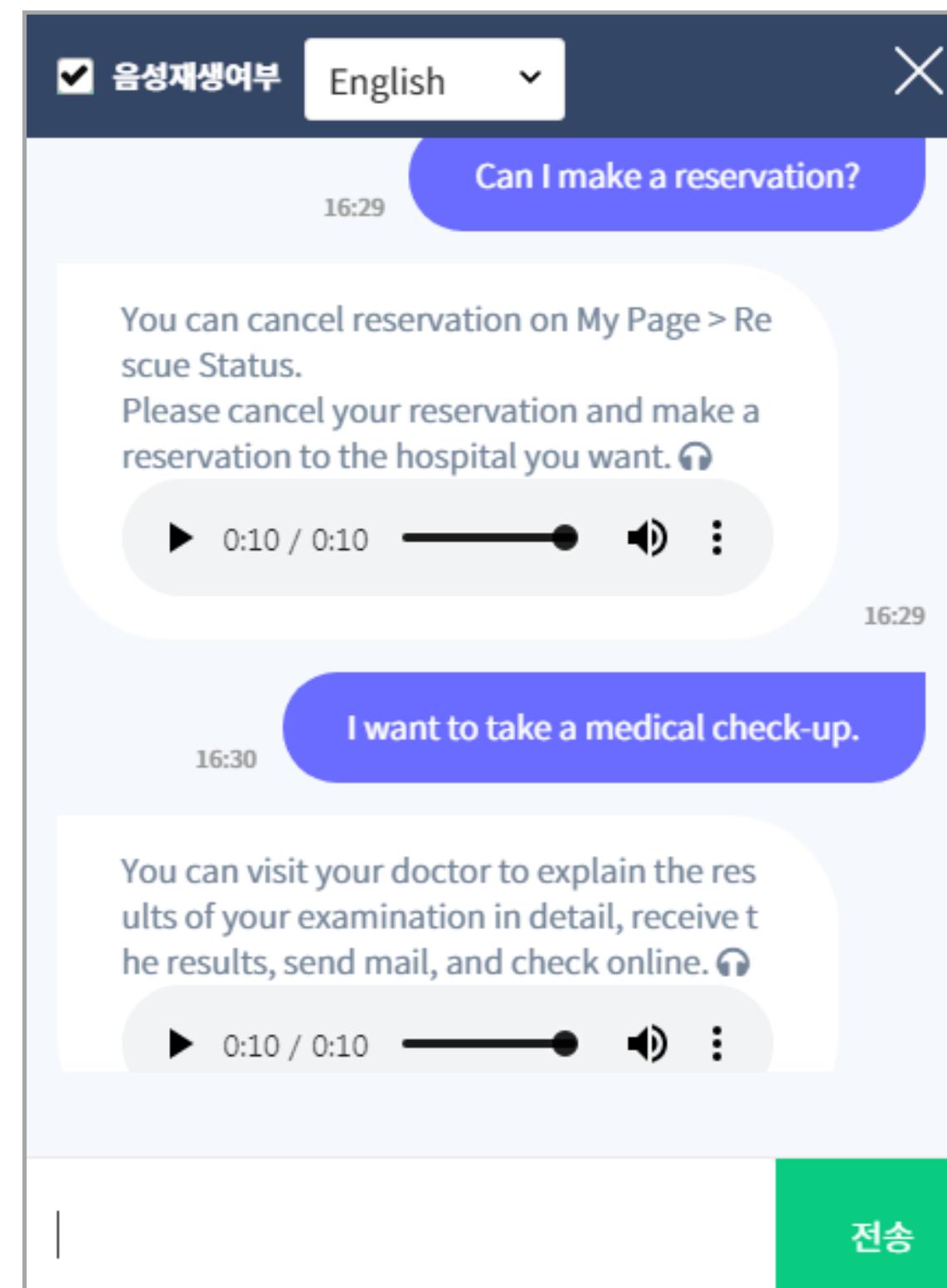
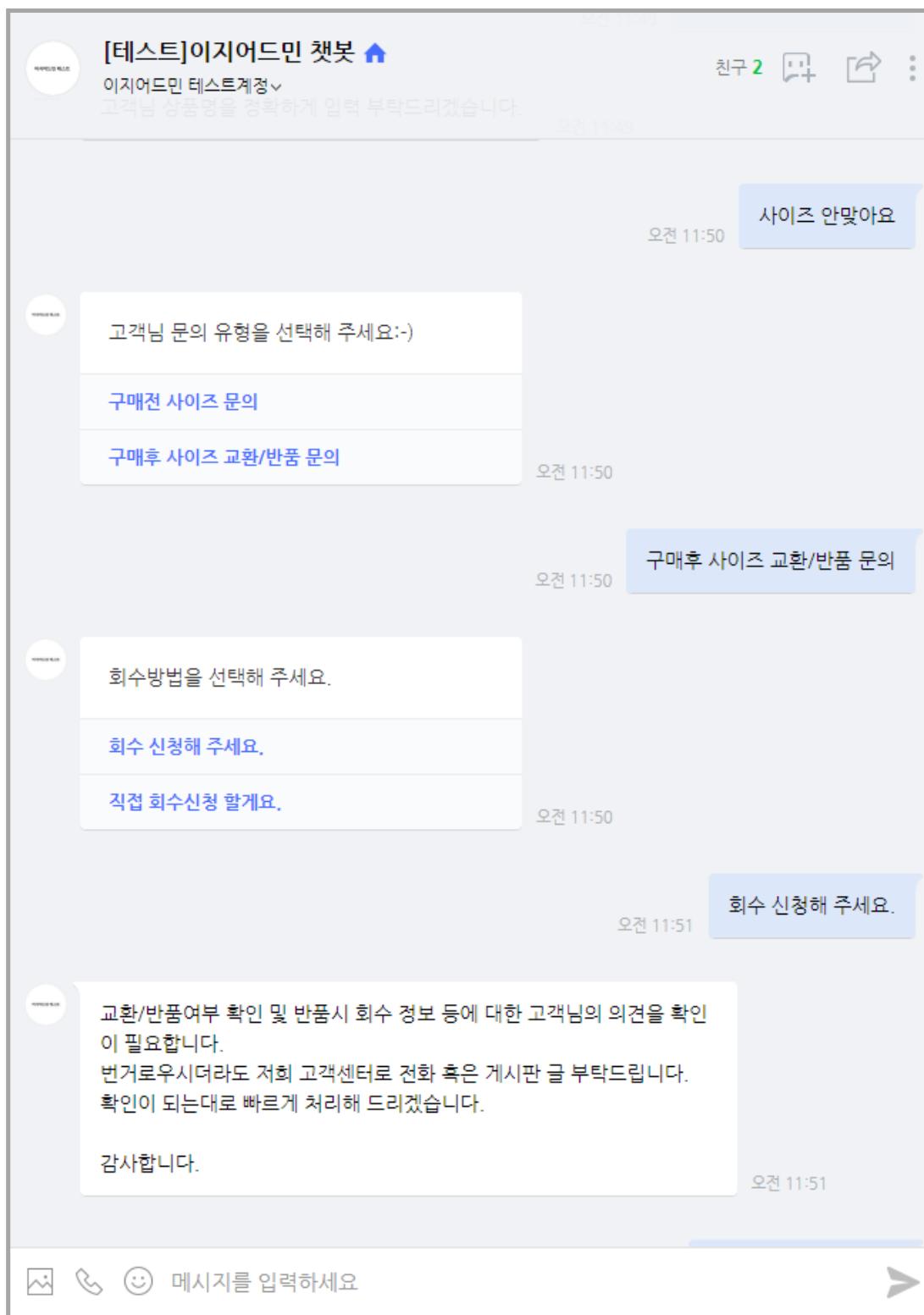
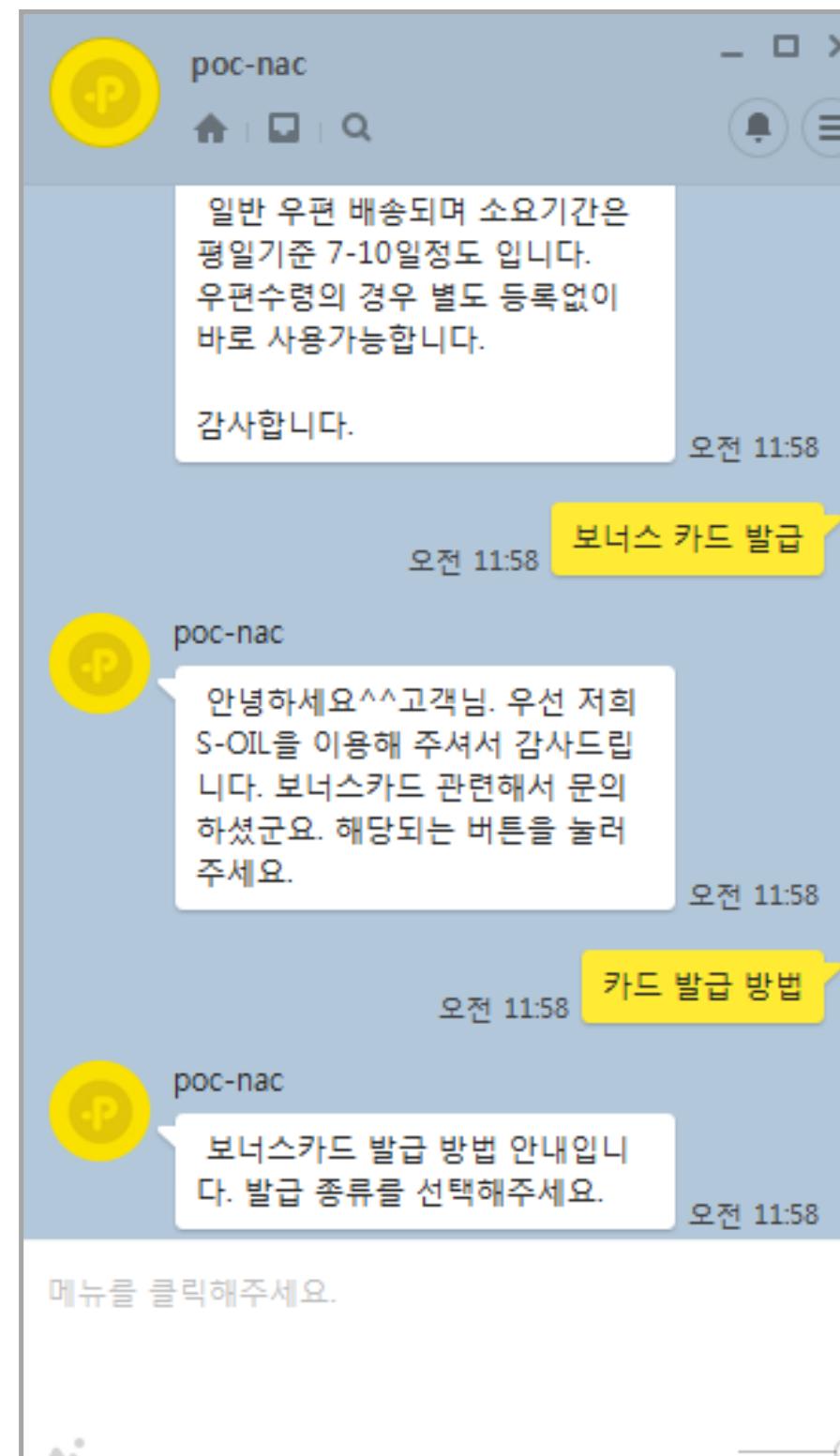


고객센터 챗봇 적용 현황 (beta)

서비스 중인 챗봇	누적 이용자 (명)	주간 이용자 (명)	품질 평가 점수*
블로그 고객센터 PC	43,451	745	83
블로그 고객센터 모바일	15,691	341	83.5
카페 고객센터 PC	14,993	903	83
카페 고객센터 모바일	12,470	837	81
지식인 고객센터 PC	467	19	82.5
지식인 고객센터 모바일	2,542	769	83
Modoo 챗봇	5,130	314	85
지식인 심심해요 봇	44,425	1,784	81.5
클로바 음성 고객센터			-

* 품질 평가 점수 : 답변을 4점 척도 기준으로 채점 (0~4점)

6.4 실제 서비스 현황 (3/3)



S-oil 고객센터 (비즈톡)

쇼핑 (핀즈)

의료 상담 (가온넷)

6.3 Closing Comments

기본 모델, 기본 아키텍쳐로도 누구나 쉽게 우수한 성능의 챗봇 제작 가능

언어별 좋은 feature를 찾기 위한 요소를 찾아보세요

단일 모델보다 다수의 baseline 모델을 ensemble 해보세요

결국 많은 실험이 필요합니다. 하지만 쉽게 테스트할 수 있는 구조가 있으면 좋겠죠?

딥러닝이 답은 아니지만, baseline model은 좋은 tool이자, 한 발 앞선 출발점입니다.

완벽한 AutoML 적용을 위해 끊임없이 노력하겠습니다.

Q & A

질문은 Slido에 남겨주세요.

sli.do

#devview

TRACK 1

Appendix

<http://www.aclweb.org/anthology/P15-1162>

<https://arxiv.org/pdf/1706.01427.pdf>

<http://www.aclweb.org/anthology/D14-1181>

<https://papers.nips.cc/paper/7181-attention-is-all-you-need.pdf>

<https://arxiv.org/pdf/1409.3215.pdf>

Thank you

