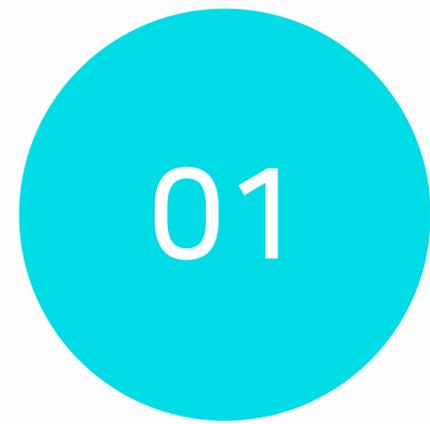


# UNICON:

## LABEL 없이 고객문의 유형을 분석 및 설계해보자

고동현 NAVER CLOVA Conversation  
최성재 NAVER CLOVA Conversation  
박정수 NAVER CLOVA Conversation

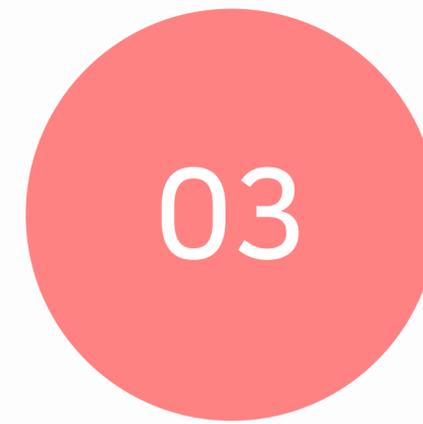
# CONTENTS



시장 분석



UNICON



SMART-FAQ



“우리가 이러한 연구 및 개발을 진행한 **이유?**”

# 1.1 커져가는 자동화 고객응대 서비스

## "스타트업 고민인 '고객응대' 믿고 맡기세요"

박민영 TWC 대표 인터뷰

문자상담·게시판 관리·콜센터  
국내 유일 고객응대 플랫폼  
'클라우드게이트'로 170社 확보

네이버 특톡·카톡상담 처리  
누적 400억원 투자 유치 성공  
"한국의 세일즈포스 될 것"

정지성 기자 | 입력 : 2021.09.22 16:03:02

## KT 고객센터, 24시간 상담 'AI 보이스봇' 전면 도입

KT 고객센터 24시간 전 분야 상담 가능  
ARS 기다림 없이 AI 보이스봇 응대

김성현 기자 입력 2021-08-18 11:33 | 수정 2021-08-18 11:33



## T전화기 AI를 만났다...SKT, 지능형전화 서비스 출시

입력 2020-10-12 17:31:01 수정 2020.10.12 18:39:53 김성태 기자



AI스피커 '누구'와 T전화 결합  
음성만으로 전화·문자 수발신  
스마트홈제어·음악 재생 등  
'누구'제공 30여 서비스도 가능  
맞춤서비스 추천 '투데이'도 탑재  
"AI 비즈플랫폼으로 발전시킬 것"

## 은행 AI 똑똑해졌네...복잡한 대출상담 척척

진화하는 신한은행 AI '솔리'

고객응성 듣고 감정상태 예측  
콜센터 문의 50% AI가 담당  
청약가입 상품도 자세히 설명  
진옥동 행장 "AI가 상담 미래"

김혜순 기자 | 입력 : 2021.09.02 17:29:39 | 수정 : 2021.09.02 20:36:28

### 진화하는 인공지능 쏘리

도입(2019년 11월)

고객 대기시간 안내 등 단순 업무

1단계 '초기형 AI'(2019년 11월~2020년 5월)

예·적금 만기일 고지 등 고객에게 전화를 거는  
통지성 업무 중 95%를 처리

2단계 '대화형 AI'(2020년 11월~2021년 7월)

고객 문의전화 50% 상담, 이 중 25%는  
전문상담사 연결 없이 자체 해결

3단계 '공감형 AI'(연내 시작목표)

AI 상담 완결률 40%로 확대, 고객 감정 인식·  
성향 분석·사물인터넷(IoT) 등으로 확장



## [VC가 짚한 스타트업] 메신저로 맞춤 상담...고객 80%가 중소기업

입력 2021-09-22 17:22:56 수정 2021.09.22 21:20:31 정다운 기자

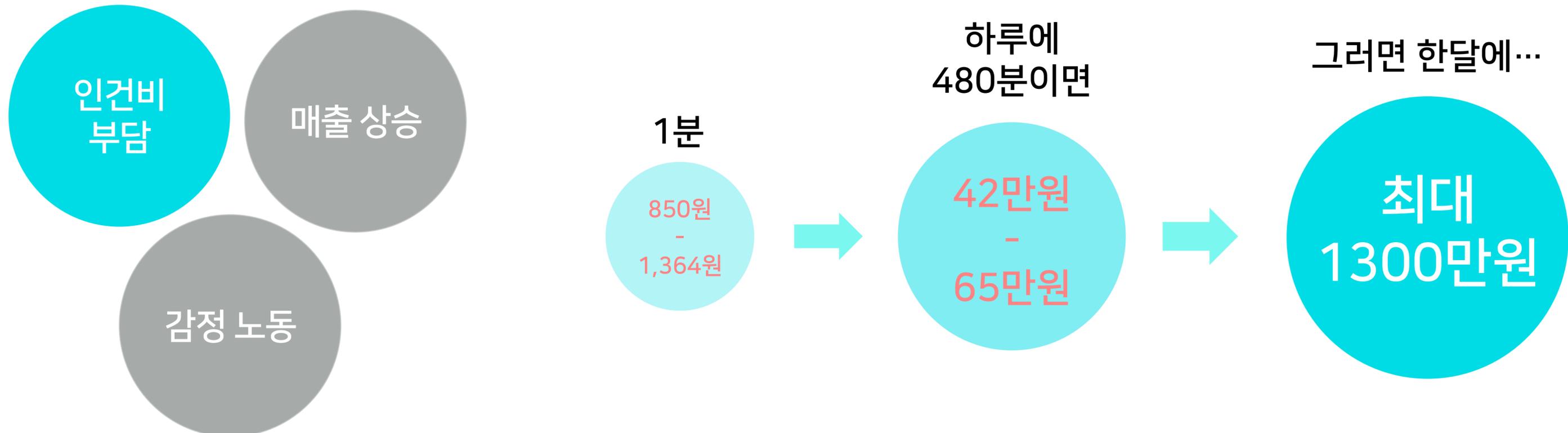


채널톡'서 고객 요구 즉각 대응  
메신저로 상담...고객 80%가 중소기업  
3년만에 고객사 6만곳 넘어서



“왜 이렇게 다들 인공지능 고객센터를 도입하려고 할까?”

# 1.2 고객 응대 시를 도입하려는 이유

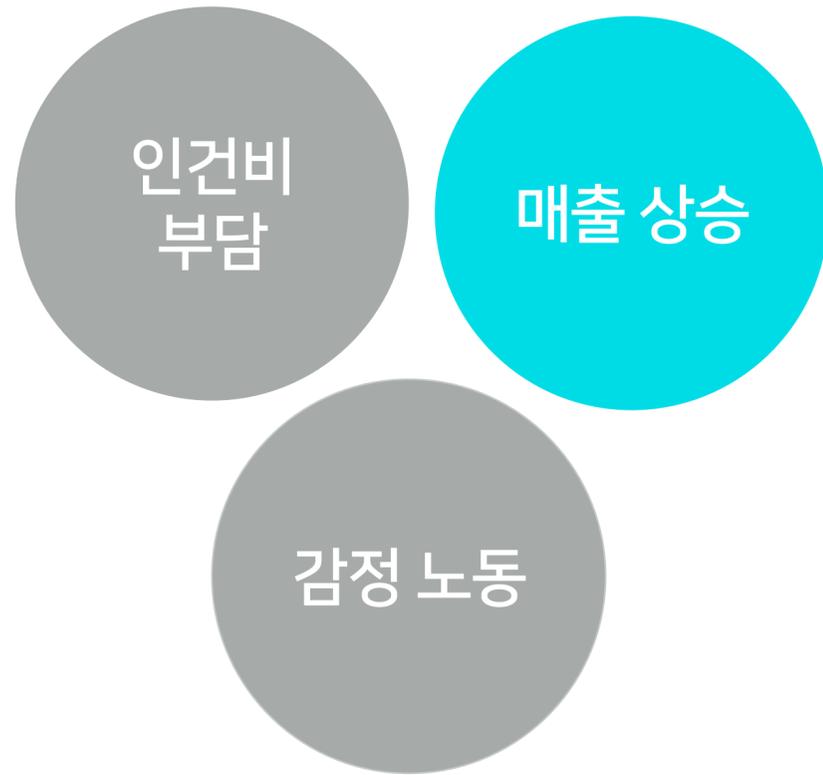


네이버 스마트 스토어 or 개인 쇼핑몰 같은 SME는

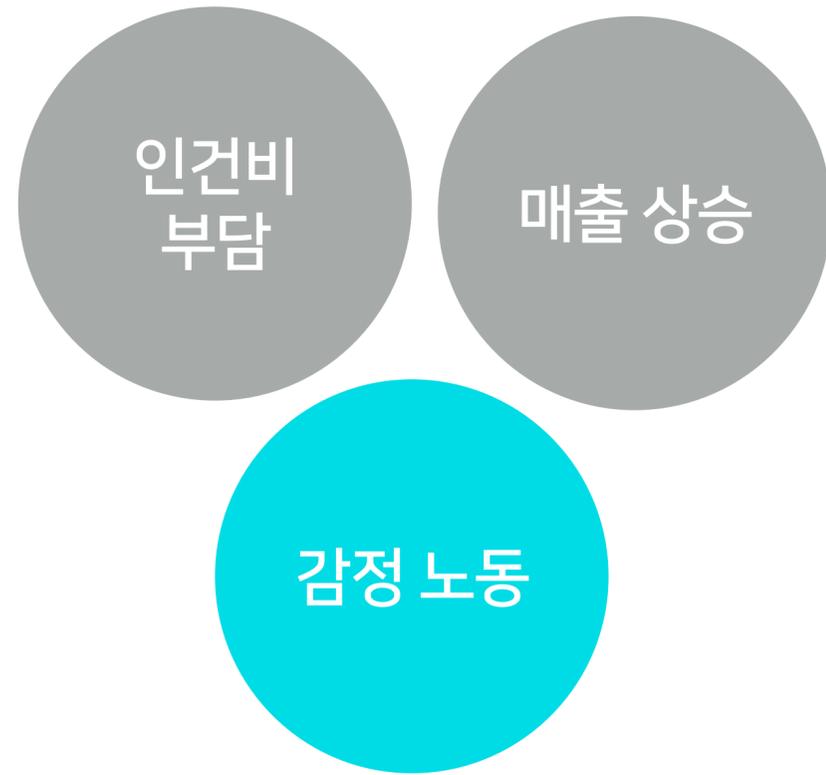
굉장히 부담!!



# 1.2 고객 응대 시를 도입하려는 이유



# 1.2 고객 응대 시를 도입하려는 이유



😓 감정노동이 많은 직업 상위 10위

- 1 텔레마케터(전화통신판매원)
- 2 호텔관리자
- 3 네일아티스트
- ⋮

고객상담까지 직접 해야 하는 1인 스토어는 ...





“인공지능 고객센터는 **어떻게** 만들어지나?”

# 1.3 인공지능고객센터는 어떻게 만들어지나?

수없이 많은 로그 데이터, 자주 문의하는 유형 분류 및 설계부터

감사합니다

이거 반품하려고 하는데요?

평균적으로 배송 며칠 걸리나요?

알겠습니다~

물건이 잘 못 온 것 같아요 반품하고 환불해주세요

어제 주문했는데 배송은 언제 올까요?

포장은 어떻게 해주시나요?

...



# 1.3 인공지능고객센터는 어떻게 만들어지나?

수없이 많은 로그 데이터, 자주 문의하는 유형 분류 및 설계부터

감사합니다

이거 반품하려고 하는데요?

평균적으로 배송 며칠 걸리나요?

알겠습니다~

물건이 잘 못 온 것 같아요 반품하고 환불해주세요

어제 주문했는데 배송은 언제 올까요?

포장은 어떻게 해주시나요?

로그  
데이터로부터  
Intent 파악

반품문의



# 1.3 인공지능고객센터는 어떻게 만들어지나?

수없이 많은 로그 데이터, 자주 문의하는 유형 분류 및 설계부터

감사합니다

이거 반품하려고 하는데요?

평균적으로 배송 며칠 걸리나요?

알겠습니다~

물건이 잘 못 온 것 같아요 반품하고 환불해주세요

어제 주문했는데 배송은 언제 올까요?

포장은 어떻게 해주시나요?

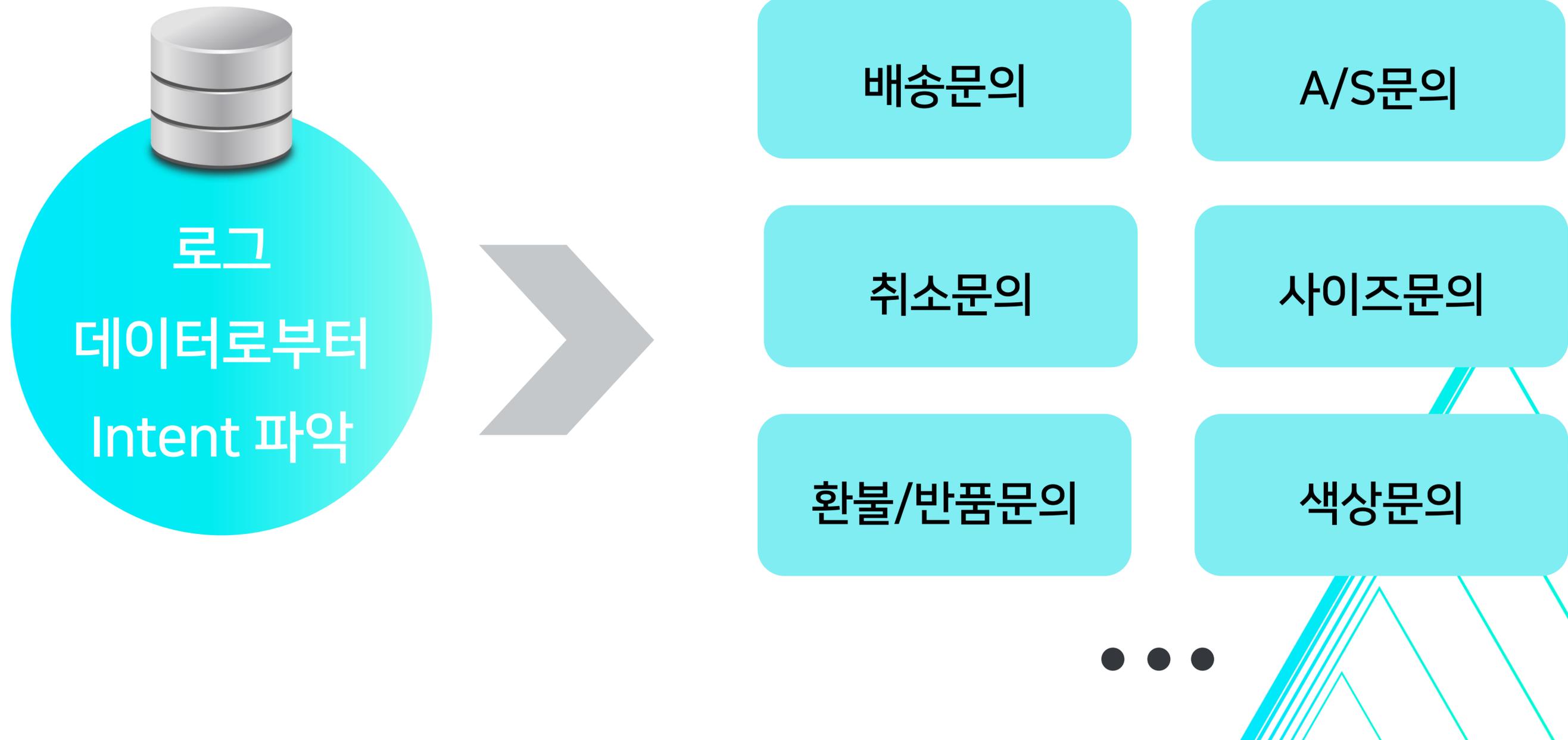
로그  
데이터로부터  
Intent 파악

배송문의



# 1.3 인공지능고객센터는 어떻게 만들어지나?

수없이 많은 로그 데이터, 자주 문의하는 유형 분류 및 설계부터



# 1.3 인공지능고객센터는 어떻게 만들어지나?

설계를 바탕으로 각 레이블에 맞는 데이터 수집



- 주문했는데 언제 오나요?
- 이번주 안에 오나요?
- 내일 발송 될까요?

배송문의

- 어제 주문했는데 취소해주세요
- 재고 없나요? 취소 부탁드립니다

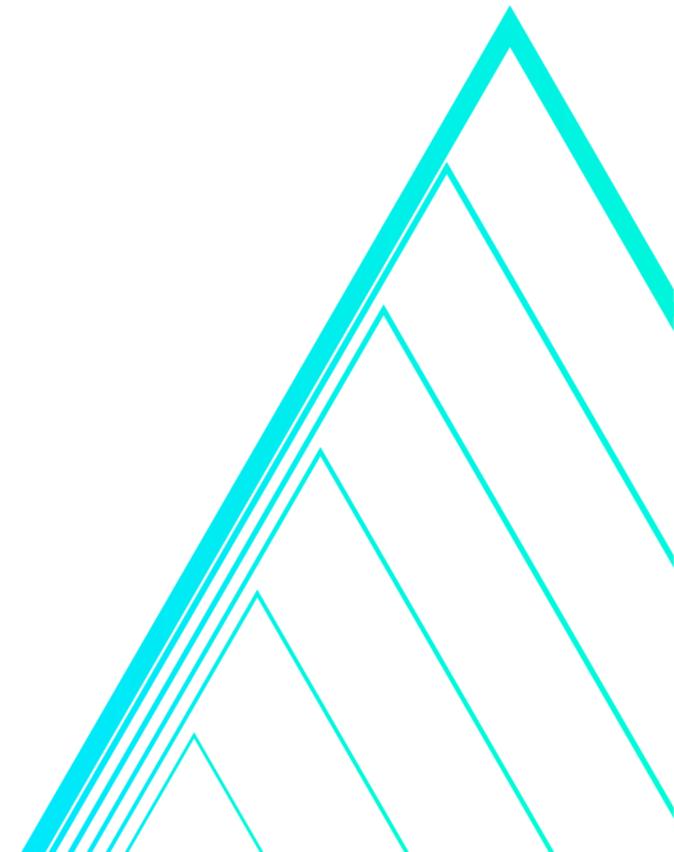
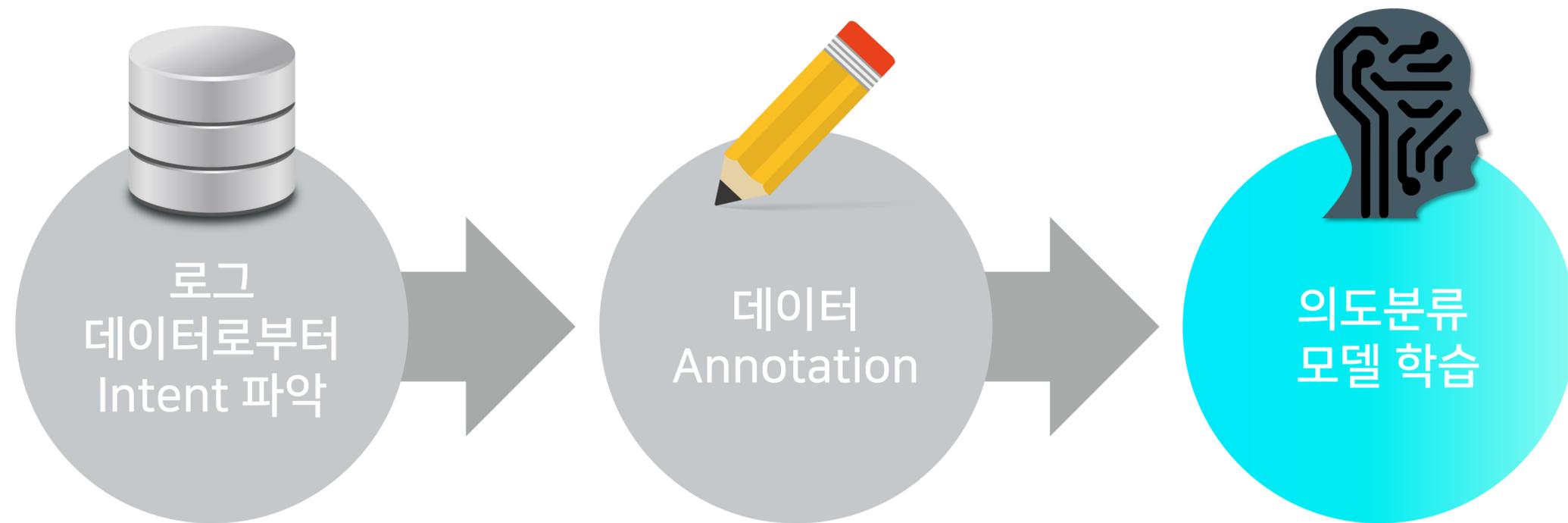
취소문의

- 제가 32정도 입는데 M이 맞을까요? L이 맞을까요?
- 제 발이 270정도인데, 발 볼 사이즈가 어떻게 나오나요?

사이즈문의

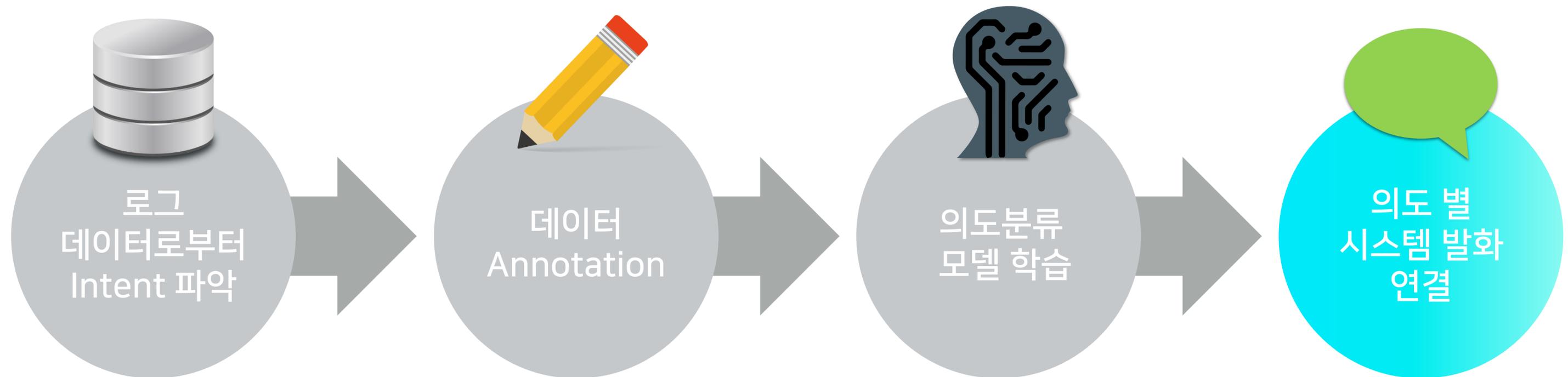
# 1.3 인공지능고객센터는 어떻게 만들어지나?

## 의도분류 모델 학습



# 1.3 인공지능고객센터는 어떻게 만들어지나?

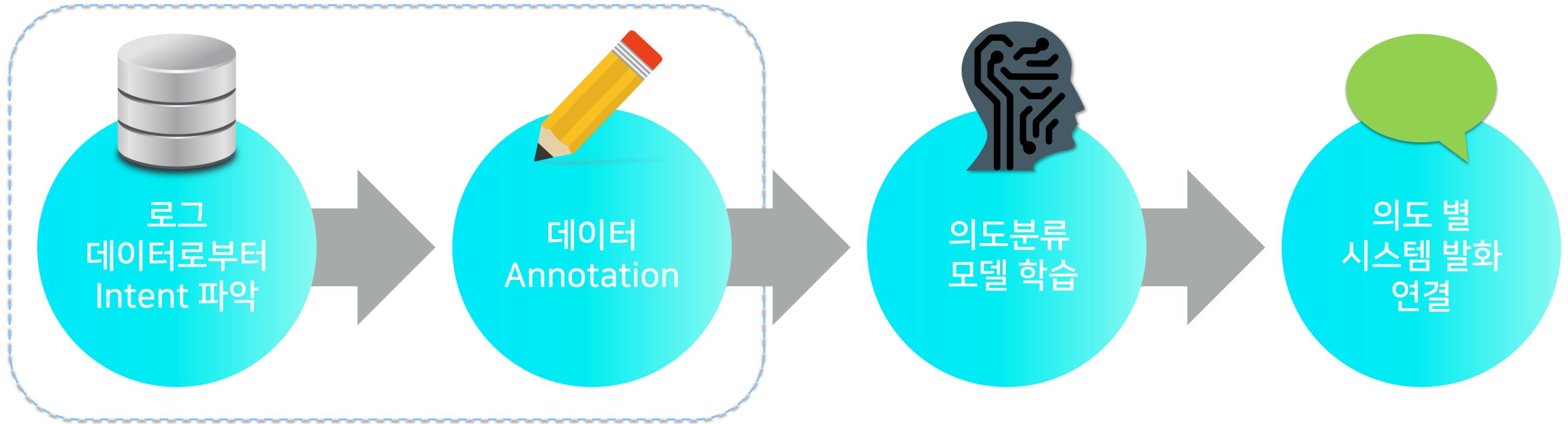
분류된 의도에 적합한 답변 연결





“하지만...”

# 1.4 인공지능고객센터 제작 비용이 문제



1. 너무 많은 데이터
2. 정의되지 않은 설계 기준
3. 크나큰 데이터셋 제작 비용

- 감당하기 쉽지 않은 **데이터 설계 비용**
- 다양한 도메인에 빠르게 **확장**이 어려움

통신, 의료, 미용실, 의류, 낚시 등

# 1.4 인공지능고객센터 제작 비용이 문제

- 1. 너무 많은 로그 데이터
- 2. 정의되지 않은 설계 기준
- 3. 크나큰 데이터 셋 제작 비용
- 4. 어려운 도메인 확장

이거 반품하려고 하는데요?      감사합니다

어제 주문했는데 배송은 언제 올까요?

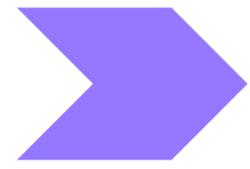
아직 주문 안들어갔으면 취소 부탁드립니다

물건이 잘 못 온 것 같아요 반품하고 환불해주세요

평균적으로 배송 며칠 걸리나요?

알겠습니다~

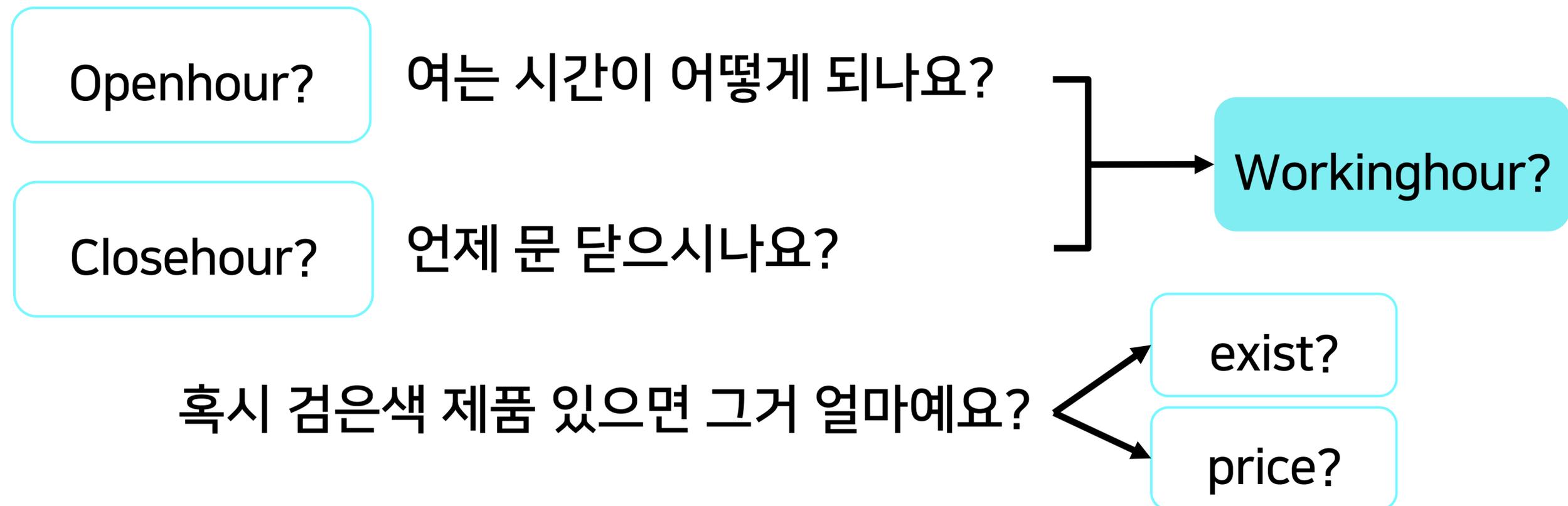
...



- 배송문의
- 취소문의
- 색상 문의
- ...

# 1.4 인공지능고객센터 제작 비용이 문제

- 1. 너무 많은 로그 데이터
- 2. 정의되지 않은 설계 기준
- 3. 크나큰 데이터 셋 제작 비용
- 4. 어려운 도메인 확장



# 1.4 인공지능고객센터 제작 비용이 문제

- 1. 너무 많은 로그 데이터
- 2. 정의되지 않은 설계 기준
- 3. 크나큰 데이터 셋 제작 비용
- 4. 어려운 도메인 확장

Dataset	# classes	# data per class
AG news	4	31,900
Dbpedia	14	45,000
IMDB	2	25,000
Average	7	40,380

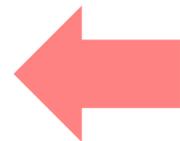
→ 도메인별 classification 모델 학습에 필요한 데이터 추정치 =  $7 * 40,380 \approx 300,000$  개

# 1.4 인공지능고객센터 제작 비용이 문제

- 1. 너무 많은 로그 데이터
- 2. 정의되지 않은 설계 기준
- 3. 크나큰 데이터 셋 제작 비용
- 4. 어려운 도메인 확장

약 삼십만 개의 데이터를 annotate하는 데에 3,500 만원 정도의 비용이 필요

Data type	Objective	Unit	Tier 1	Tier 2
Image	Classification	Image	\$35	\$25
	Bounding box	Bounding box	\$63	\$49
	Segmentation	Segment	\$870	\$850
	Rotated box	Bounding box	\$86	\$60
	Polygon/polyline	Polygon/Polyline	\$257	\$180
Video	Classification	5sec video	\$86	\$60
	Object tracking	Bounding box	\$86	\$60
	Event	Event in 30sec video	\$214	\$150
	Classification	50 words	\$129	\$90
Text	Entity extraction	Entity	\$86	\$60



# 1.4 인공지능고객센터 제작 비용이 문제

- 1. 너무 많은 로그 데이터
  - 2. 정의되지 않은 설계 기준
  - 3. 크나큰 데이터 셋 제작 비용
  - 4. 어려운 도메인 확장
- 반복

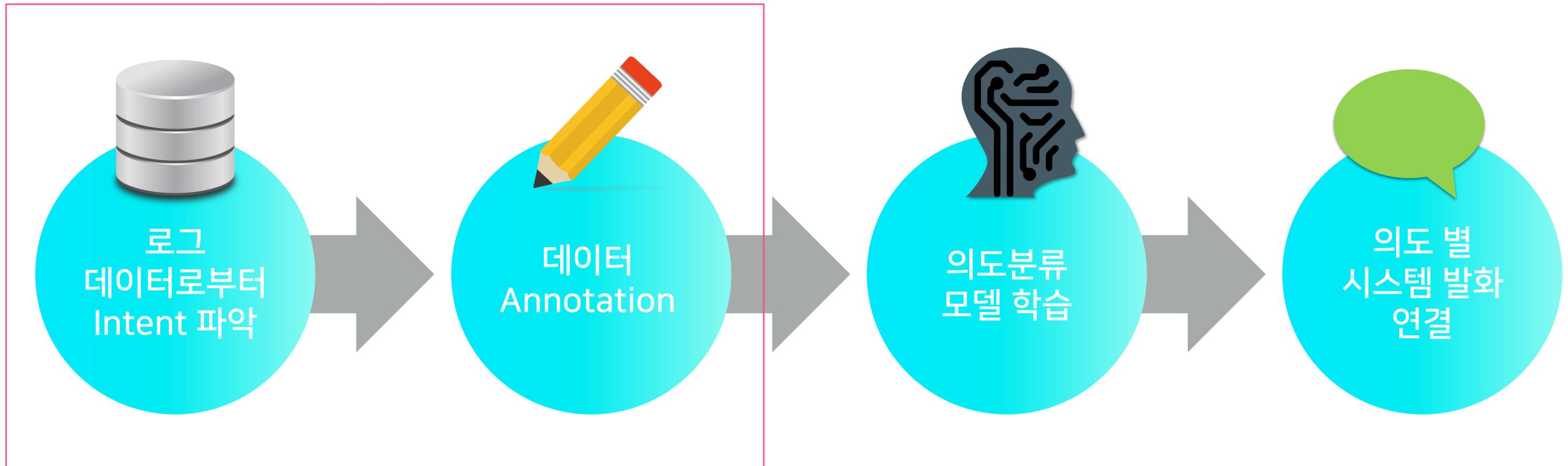
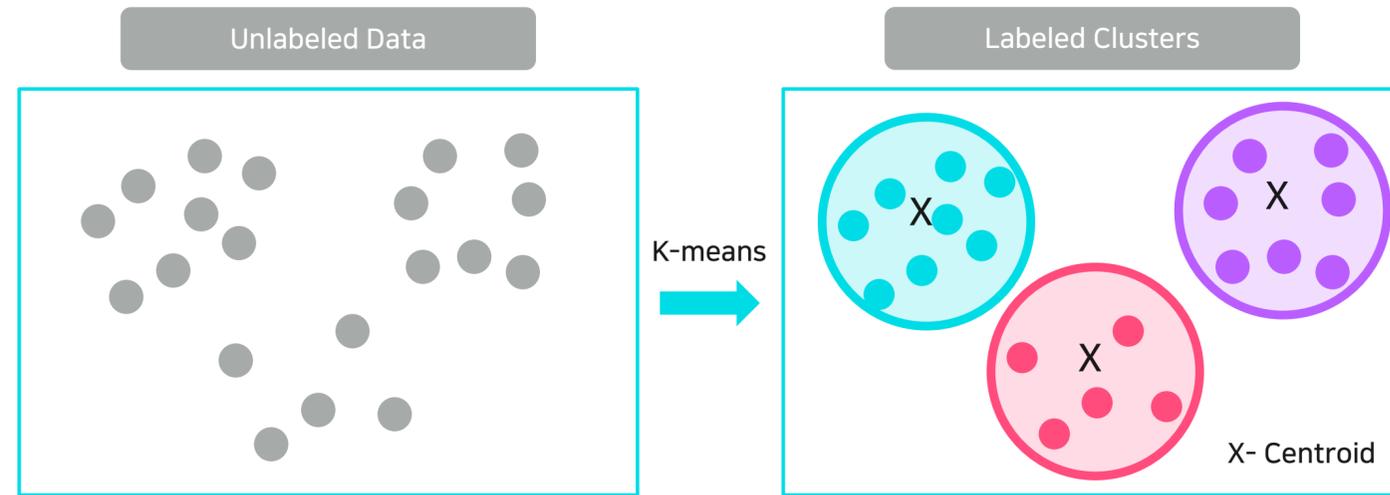
## 식당

예약되나요? 포장할 수 있나요?  
메뉴 추천 좀 해주세요  
매장에 자리 있나요?

## 의류점

제품 사이즈 있나요?  
언제쯤 배송 될까요?  
교환은 언제까지 가능한가요?

# 1.5 Clustering to the Rescue

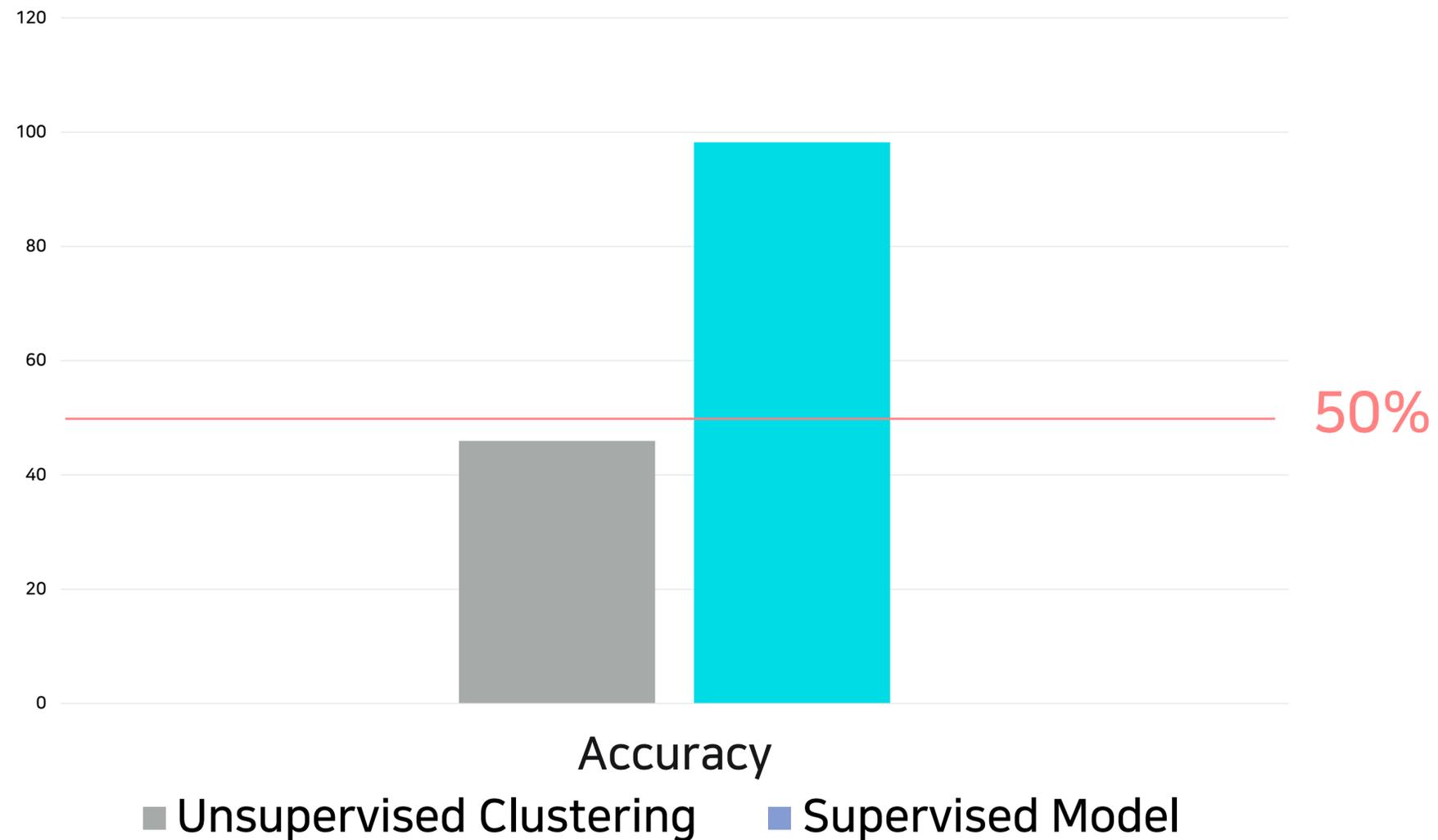


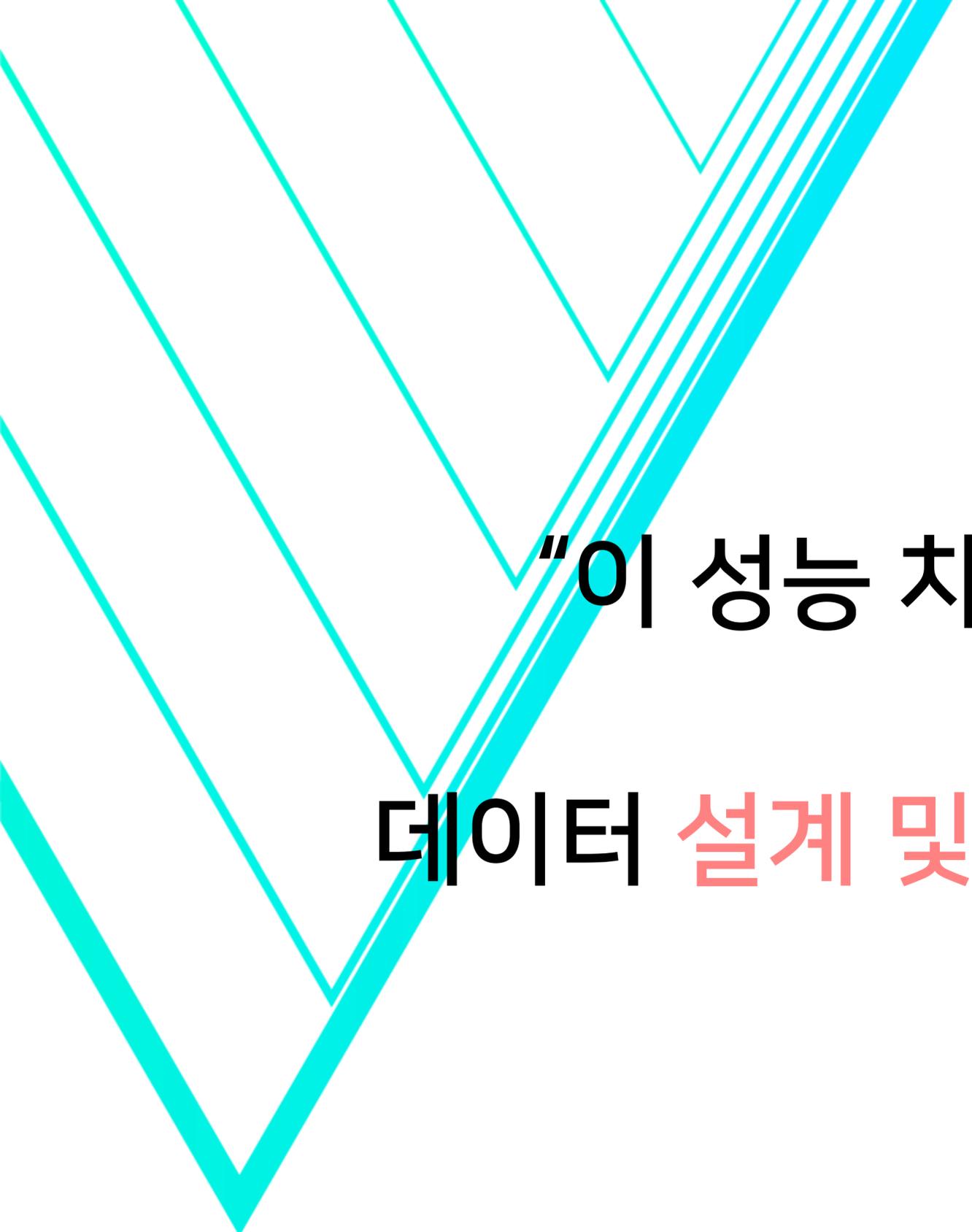


“이때까지 왜 clustering을 사용하고 있지 않았을까?”

# 1.5 Clustering to the Rescue

기존의 clustering은 label 이 없으면 50% 성능도 나오지 않는...





“이 성능 차이를 좁힐 수 있으면...”

데이터 설계 및 제작 비용을 줄일 수 있음”

# 2. UNICON

Unsupervised Intent Discovery by Contrastive Learning

## 2.1 UNICON이란?

Contrastive Learning 을 기반으로, 레이블링 된 데이터 없이

의미적으로 유사한 문장들을

Embedding Space 상에서 가깝게 위치하도록 학습하여

Classification & Clustering 등의 Task에서 사용할 수 있도록 하는 Model

## 2.1 UNICON이란?

Contrastive Learning 을 기반으로, 레이블링 된 데이터 없이

의미적으로 유사한 문장들을

Embedding Space 상에서 가깝게 위치하도록 학습하여

**Better and Efficient Sentence Representation**

Classification & Clustering 등의 Task에서 사용할 수 있도록 하는 Model

# 2.2 Sentence Representation

문장의 의미를 대표할 수 있는 real-valued representation vector

“이번에 DEVIEW 발표를 하게 된 최성재입니다” → [0.1, 0.7, ..., 0.3, 0.5]



★★★★★ 10  
[관람객] 연인들을 위한 영화가 아니다. 여성관객들을 위한 영화도 아니다. 이 영화는 꿈꾸는 모든 이들을 위한 영화. 인생을 논하는 영화.  
dhff\*\*\*\* | 2016.12.08. 16:27 | 신고  
6343 305

긍정

★☆☆☆☆ 1  
[관람객] 어... 왜 이게 인생영화라는건지 1도 모르겠어영π 노래도 좋긴한데 전체적으로 좀 지겨운 느낌..  
skye\*\*\*\* | 2016.12.19. 19:16 | 신고  
544 441

부정



Sentiment Classification

Information Retrieval

# 2.2 Sentence Representation

1954 - 2012



**BoW, TF-IDF**

Term and document frequency based model

2013 - 2014



**Word2Vec, GloVe**

Co-occurrence matrix based neural network model

2018



**Transformers (BERT)**

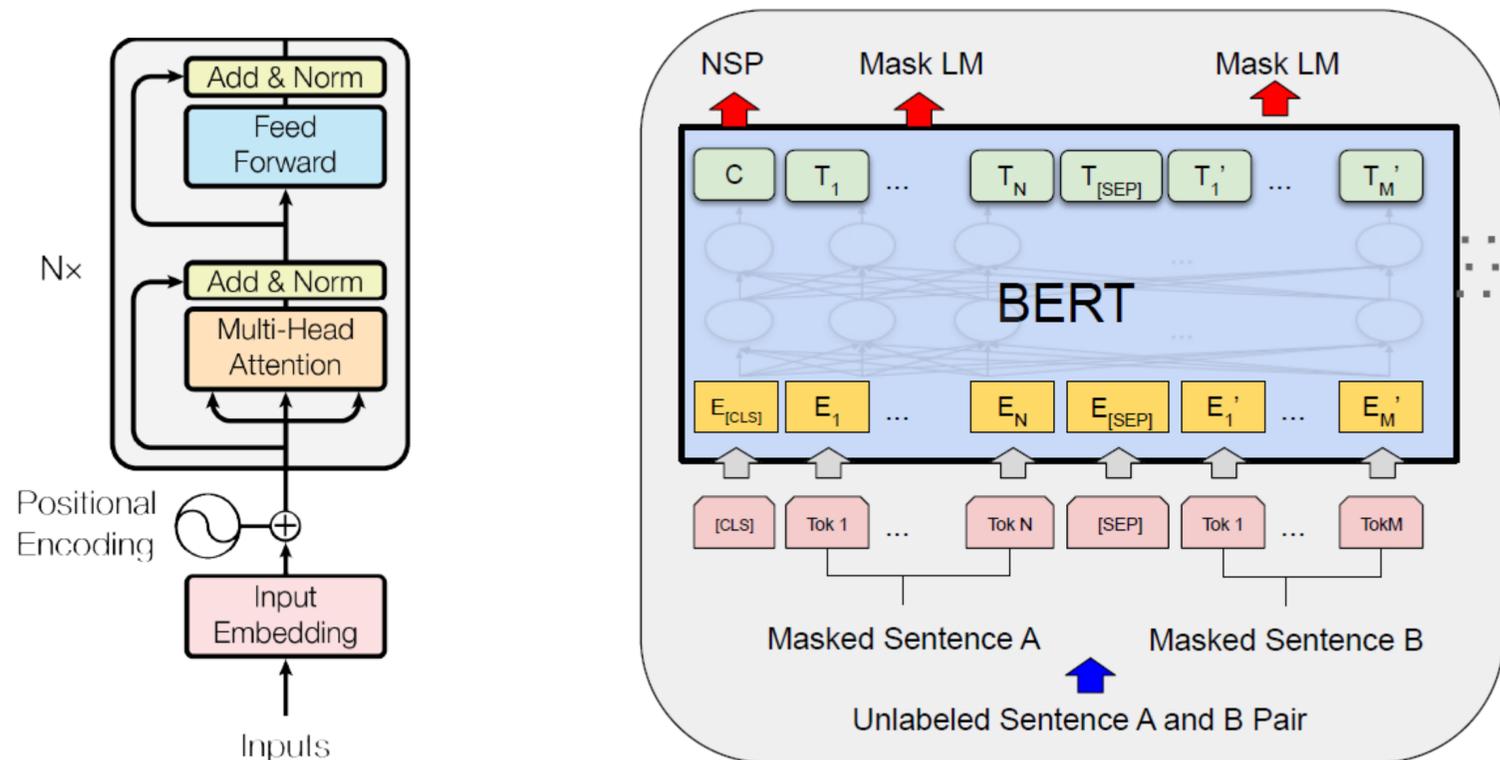
Unsupervised pre-trained language model with large-scale corpus

2019



**SBERT**

Supervised pre-trained language model with NLI datasets



BERT, RoBERTa, ALBERT and ETC.

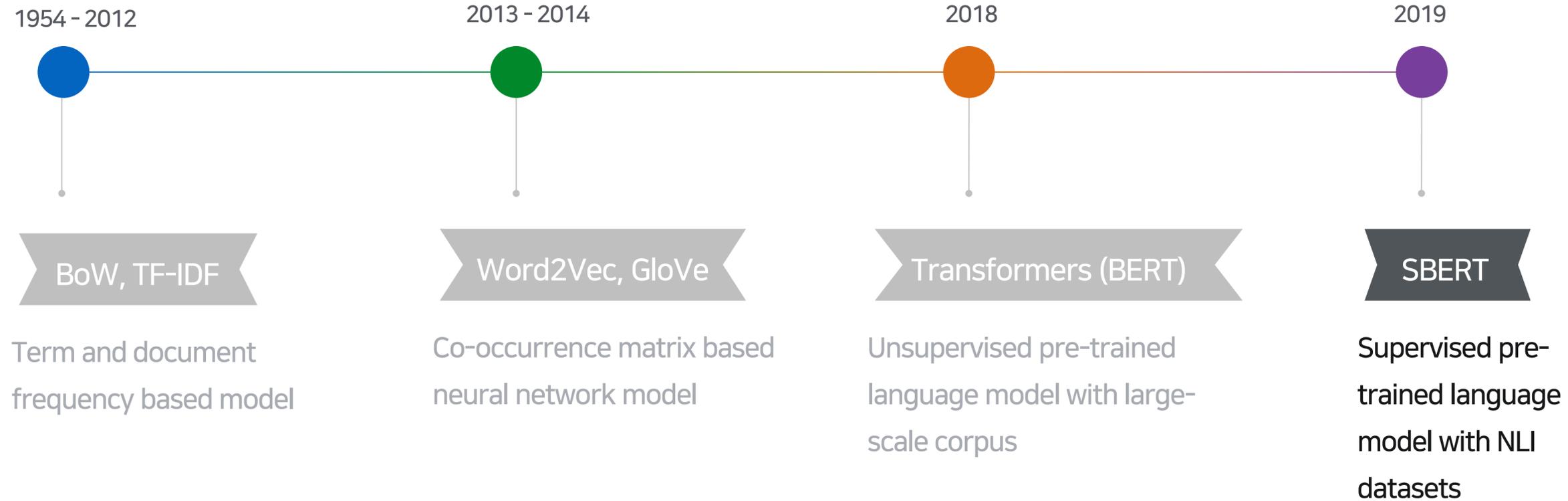
## Pros

- 문맥을 고려할 수 있음
- 다양한 downstream task 에서, 적은 양의 fine-tuning 으로도 뛰어난 성능을 얻을 수 있음

## Cons

- Fine-tuning 없이는, 서로 다른 문장 사이의 관계를 제대로 표현할 수 없음

# 2.2 Sentence Representation



Model	STS12	STS13	STS14	STS15	STS16	STSb	SICK-R	Avg.
Avg. GloVe embeddings	55.14	70.66	59.73	68.25	63.66	58.02	53.76	61.32
Avg. BERT embeddings	38.78	57.98	57.98	63.15	61.06	46.35	58.40	54.81
BERT CLS-vector	20.16	30.01	20.09	36.88	38.08	16.50	42.63	29.19
InferSent - Glove	52.86	66.75	62.15	72.77	66.87	68.03	65.65	65.01
Universal Sentence Encoder	64.49	67.80	64.61	76.83	73.18	74.92	<b>76.69</b>	71.22
SBERT-NLI-base	70.97	76.53	73.19	79.09	74.30	77.03	72.91	74.89
SBERT-NLI-large	72.27	<b>78.46</b>	<b>74.90</b>	80.99	76.25	<b>79.23</b>	73.75	76.55
SRoBERTa-NLI-base	71.54	72.49	70.80	78.74	73.69	77.77	74.46	74.21
SRoBERTa-NLI-large	<b>74.53</b>	77.00	73.18	<b>81.85</b>	<b>76.82</b>	79.10	74.29	<b>76.68</b>

## Pros

- Fine-tuning 없이도, 기존의 BERT 를 뛰어넘는 문장 유사도 성능

## Cons

- 많은 양의 Label 된 NLI 데이터셋 필요

NLI: Natural Language Inference

# 2.2 Sentence Representation

요약하자면...

Representation 을 만들기 위한 많은 방법론 들이 있지만,

우리가 원하는 것은 아니었음

	BOW	Word2Vec	BERT	SBERT
Word Similarity	✗	✓	✓	✓
Word Order & Context	✗	✗	✓	✓
W/O Labeled Data	✓	✓	✓	✗
Sentence Similarity	✓	✓	✗	✓

“우리가 원하는 것은

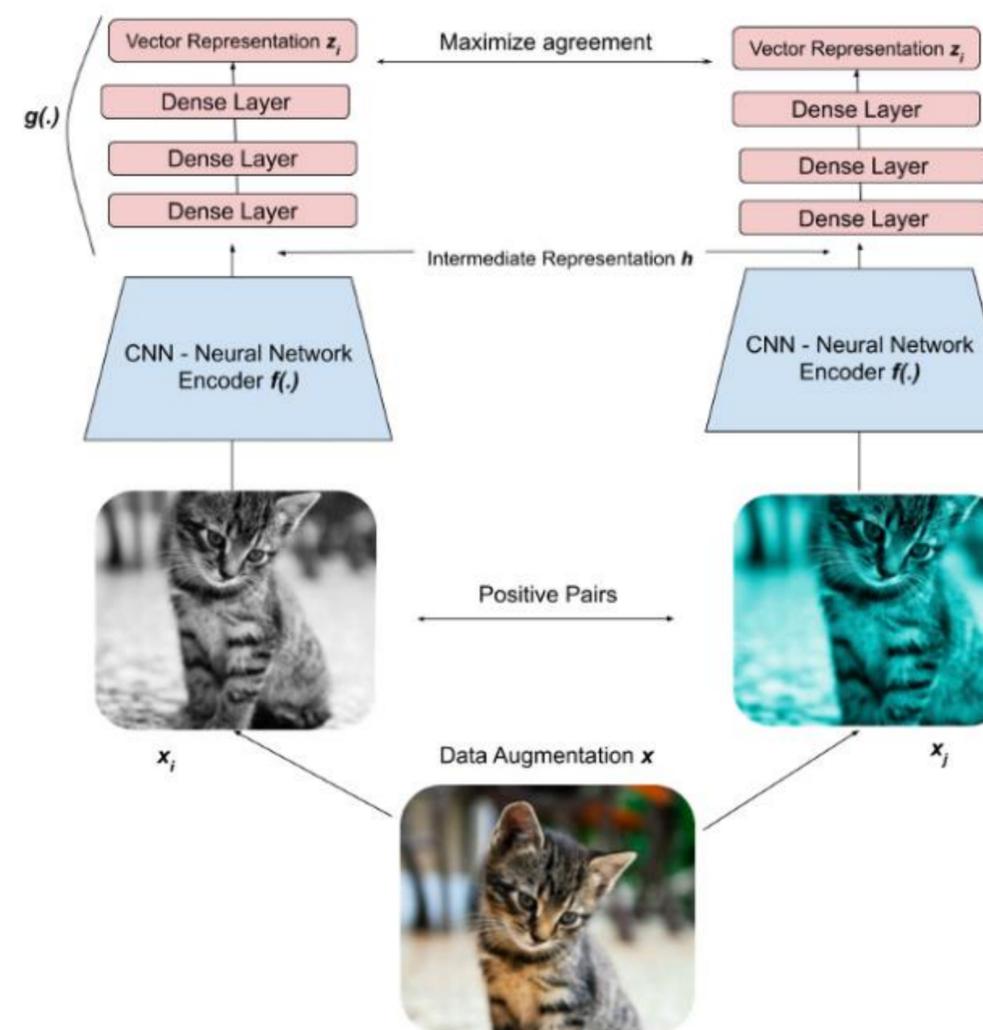
Label 된 데이터 없이

SBERT 이상의 semantic search 성능”

# 2.3 Contrastive Learning

## Contrastive Learning이 최근에 핫 하더라구요

- SimCLR, CLEAR, SimCSE, ConSERT, ETC.
- **Data augmentation** 을 통해, positive-pair 를 추출
- Positive-pair 에 대해서 **가깝게**, 그 외의 데이터에 대해 **멀리** 위치하도록 학습
- **Label 없이도 SBERT 와 유사한 semantic search 성능**



An overview of the SimCLRv2 framework. Figure by Author. Photo by [Edgar](#) on [Unsplash](#)

“하지만 Contrastive Learning도 문제가 있으니...”

## 2.3 Contrastive Learning

### 1. Data Augmentation 을 하면 Label 자체가 바뀔 수 있어요

동의어 사전

- WordNet Augmenter
- PLM based Contextual Augmenter
- Paraphrase via back translation



Deletion / Insertion / Substitution

└─┬─> Generation

## 2.3 Contrastive Learning

### 1. Data Augmentation 을 하면 Label 자체가 바뀔 수 있어요

- Intent Detection Dataset 에 대해 학습하고,
- Original test set 과 Augmented test set 에 대해 평가

Original Test Set

95.02%

Augmented Test Set

84.08%

10.94% 하락

- BERT/RobERTa 의 BLEU score 를 이용한 word substitute data augmentation 활용
- <https://github.com/makcedward/nlpaug>

## 2.3 Contrastive Learning

### 1. Data Augmentation 을 하면 Label 자체가 바뀔 수 있어요

please **remind** me at a later

[ reminder\_update ]



please **kiss** me at later time

[ goodbye ]

Find a good **recipe** for meatloaf

[ recipe ]



Find a good **flavor** for meatloaf

[ ingredients\_list ]

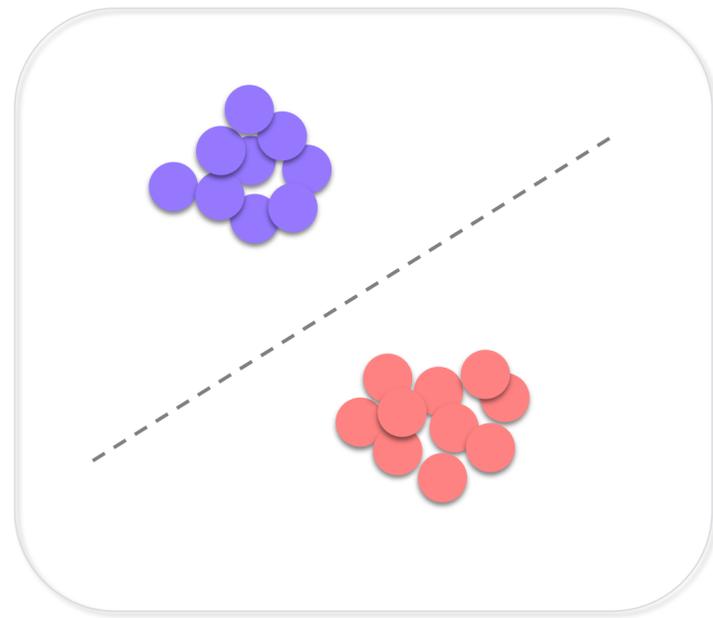
*“Data Augmentation의 문제점을 해결하기 위해  
많은 분들이 연구하고 있어요”*

[DEVIEW2021] 세상에 쓸모없는 데이터는 없다: HyperCLOVA를 이용한 반지도 학습

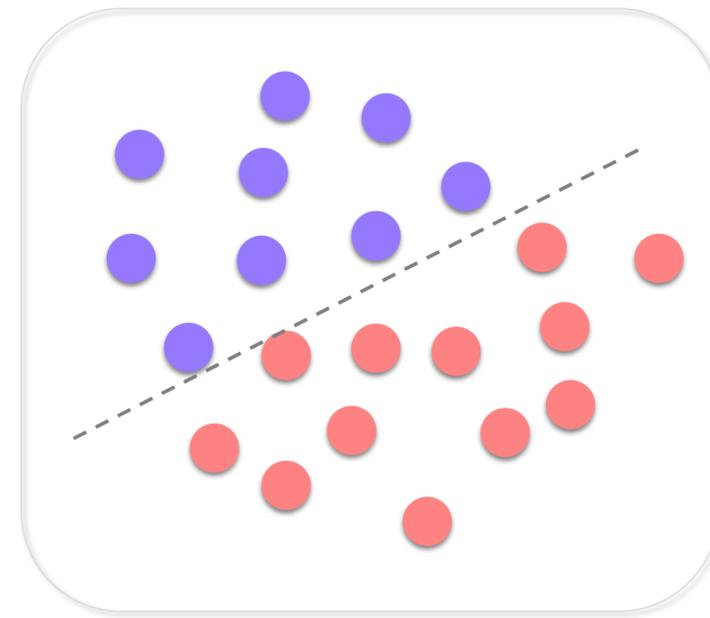
유강민, 박동주 @ NAVER / Conversation

## 2.3 Contrastive Learning

1. Data Augmentation 을 하면 Label 자체가 바뀔 수 있어요
2. Clustering에 적합한 형태가 되지 않아요



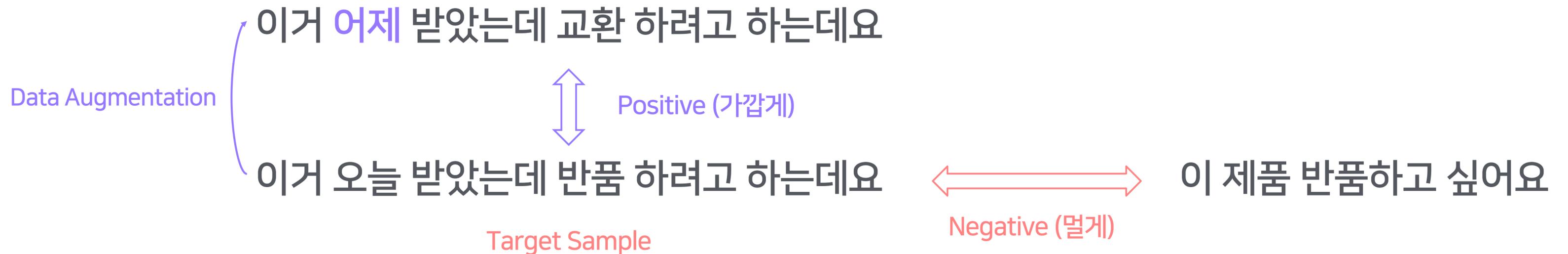
우리가 원하는 것



현실

## 2.3 Contrastive Learning

1. Data Augmentation 을 하면 Label 자체가 바뀔 수 있어요
2. Clustering에 적합한 형태가 되지 않아요



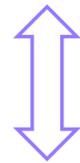
## 2.3 Contrastive Learning

1. Data Augmentation 을 하면 Label 자체가 바뀔 수 있어요
2. Clustering에 적합한 형태가 되지 않아요

극단적으로는...

Data Augmentation

이거 오늘 받았는데 교환 하려고 하는데요



Positive (가깝게)

이거 오늘 받았는데 반품 하려고 하는데요

Target Sample



Negative (멀게)

이 제품 반품하고 싶어요

이 문제를 해결하고자,

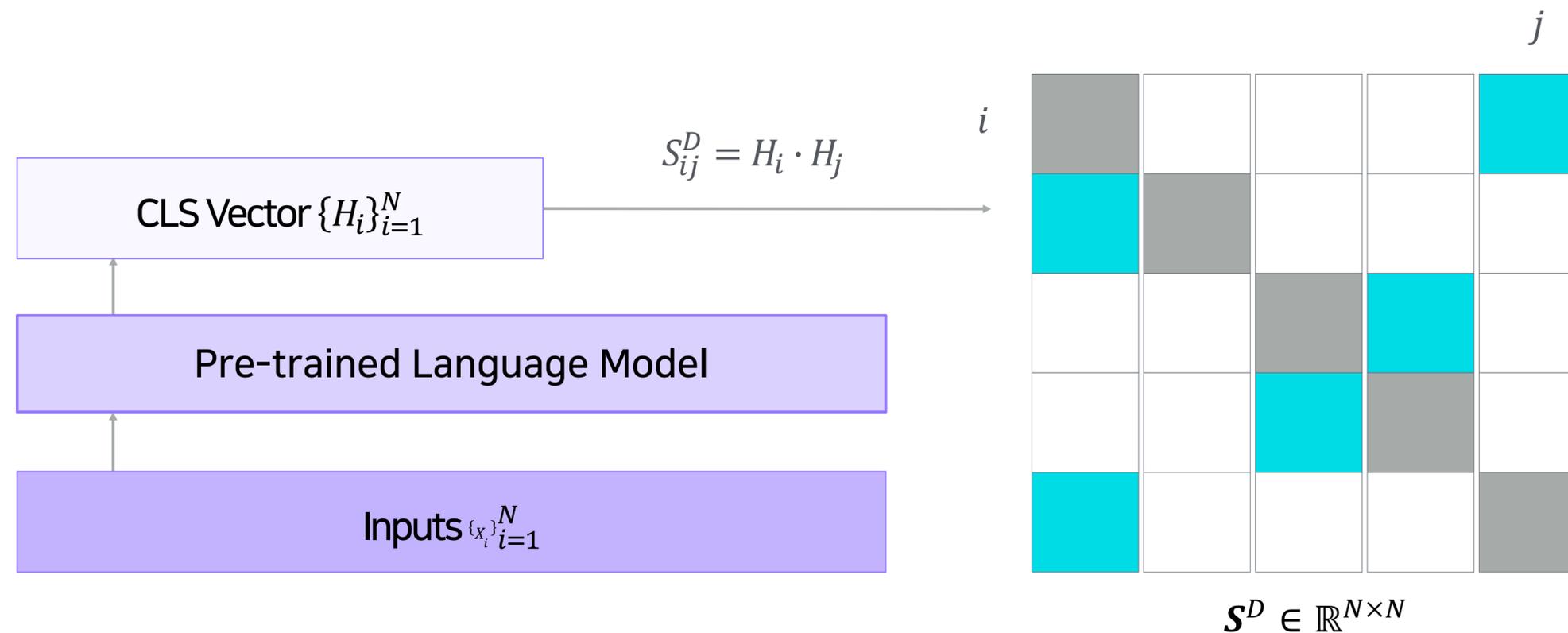
*“Data Augmentation 없이  
Clustering에 적합한”*

Contrastive Learning을 사용하는 UNICON 을 개발

## 2.4 UNICON의 구조

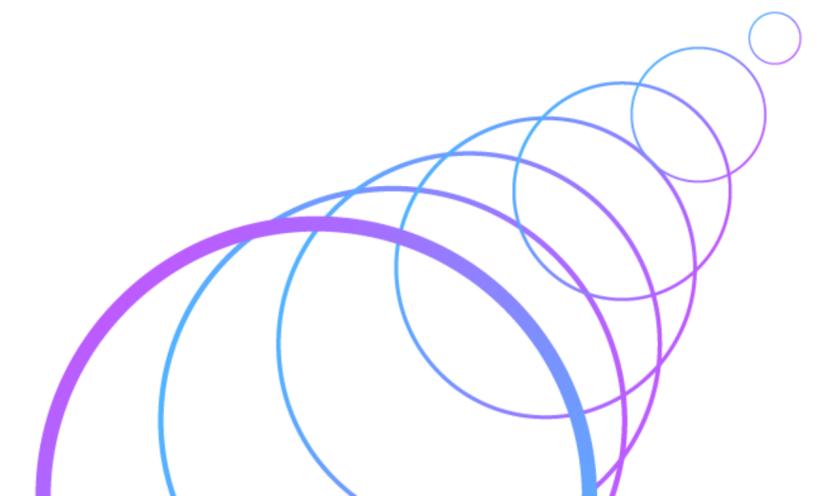
### Pseudo-label Contrastive Learning

- Data Augmentation을 활용하지 않고, positive pair를 추출하기 위해
- Batch 내의 문장 간에 유사도를 측정하여, 유사도가 가장 큰 문장을 positive sample 이라 하고 나머지를 negative samples라 정의



$$L_{ij} = \frac{\exp(\text{sim}(H_i, H_j)/\tau)}{\sum_{k=1, [k \neq i]}^N \exp(\text{sim}(H_i, H_k)/\tau)}$$

NT-Xent Loss

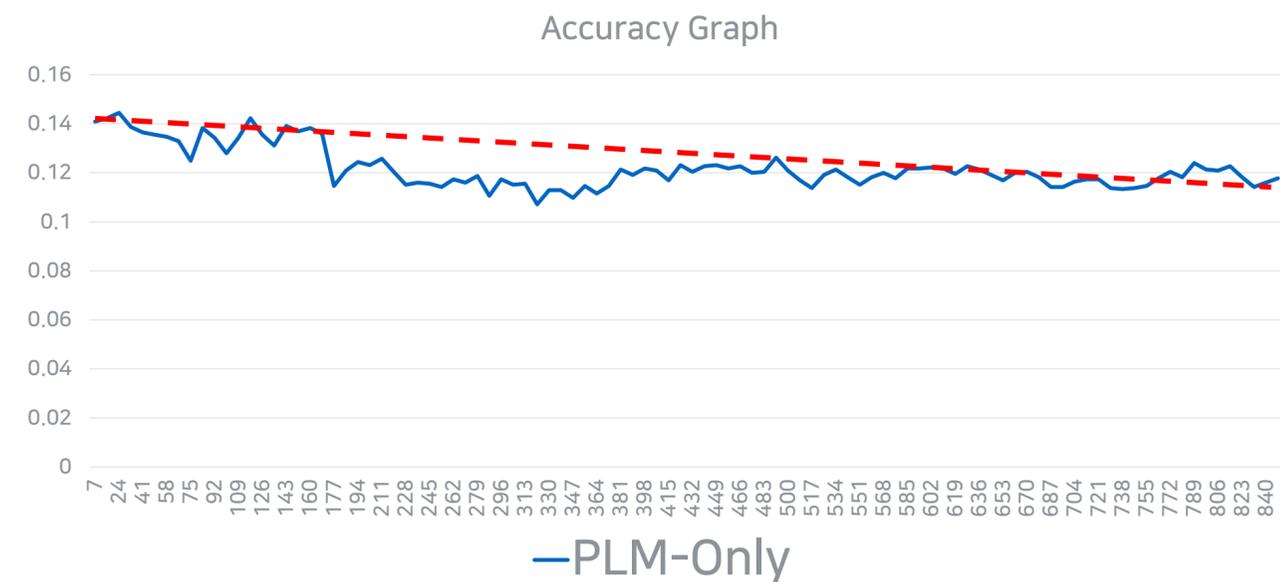
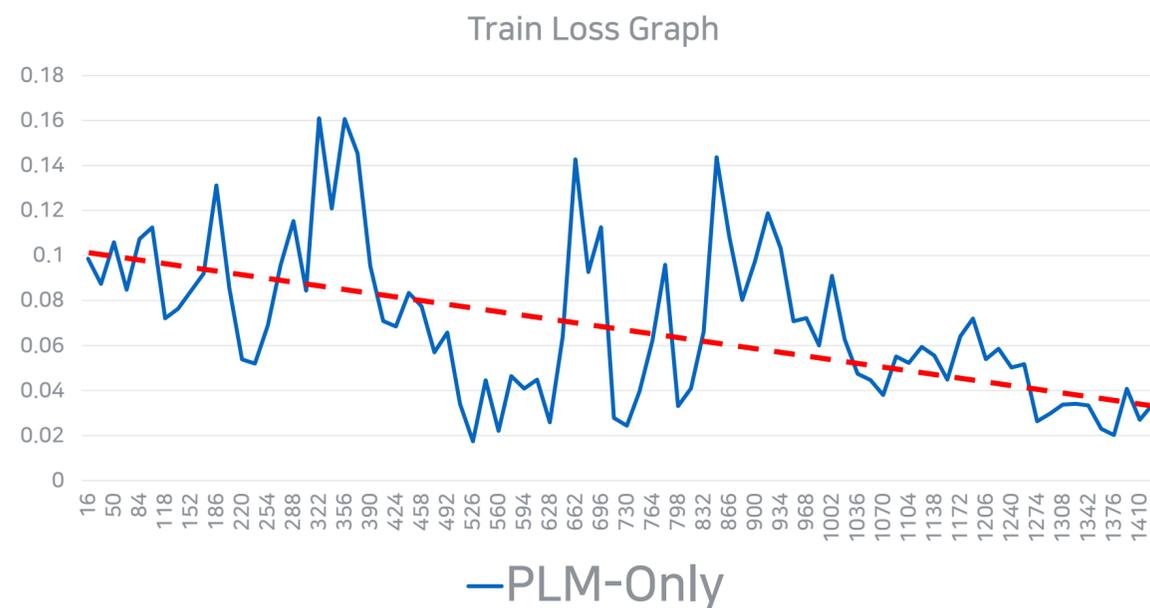


## 2.4 UNICON의 구조

두 데이터 간의 유사도를 이용해 positive-pair 를 찾아보자

Model	STS12	STS13	STS14	STS15	STS16	STSB	SICK-R	Avg.
BERT CLS-vector	20.16	30.01	20.09	36.88	38.08	16.50	42.63	29.19

하지만, BERT 만 사용해서는 제대로 된 positive-pair 를 찾을 수 없었음

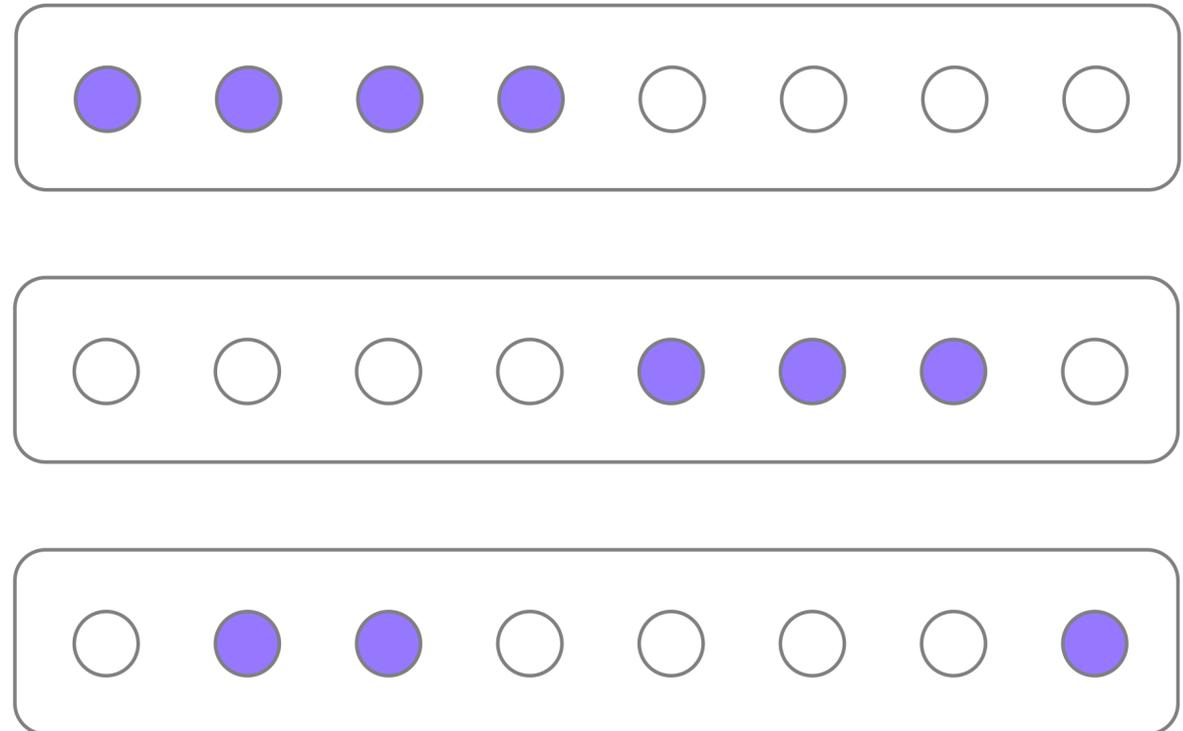


# 2.4 UNICON의 구조

## Auxiliary Representation

### 초기에 불안정한 BERT를 가이드할 수 있는 보조적인 수단을 추가하자

- 1. BOW (Bag-Of-Words)
  - 2. TF-IDF (Term Frequency-Inversed Document Frequency)
  - 3. BM25
- 1. label이 필요하지 않음
  - 2. 적당한 semantic search 성능



## 2.4 UNICON의 구조

Auxiliary Representation

# 초기에 불안정한 BERT를 가이드할 수 있는 보조적인 수단을 추가하자

4. Word2Vec (CBOW, Skip-Gram)

5. GloVe

나는 오늘 네이버 DEVIEW 에서 발표를 하고 있습니다

DEVIEW는 네이버에서 주최하는 기술 컨퍼런스입니다

왕 - 남자 + 여자 = 여왕

단어 간에 유의미한 산술 연산이 가능

$\text{sim}(\text{DEVIEW}, \text{네이버}) \gg \text{sim}(\text{DEVIEW}, \text{카카오})$

단어 간의 유사성 비교 가능

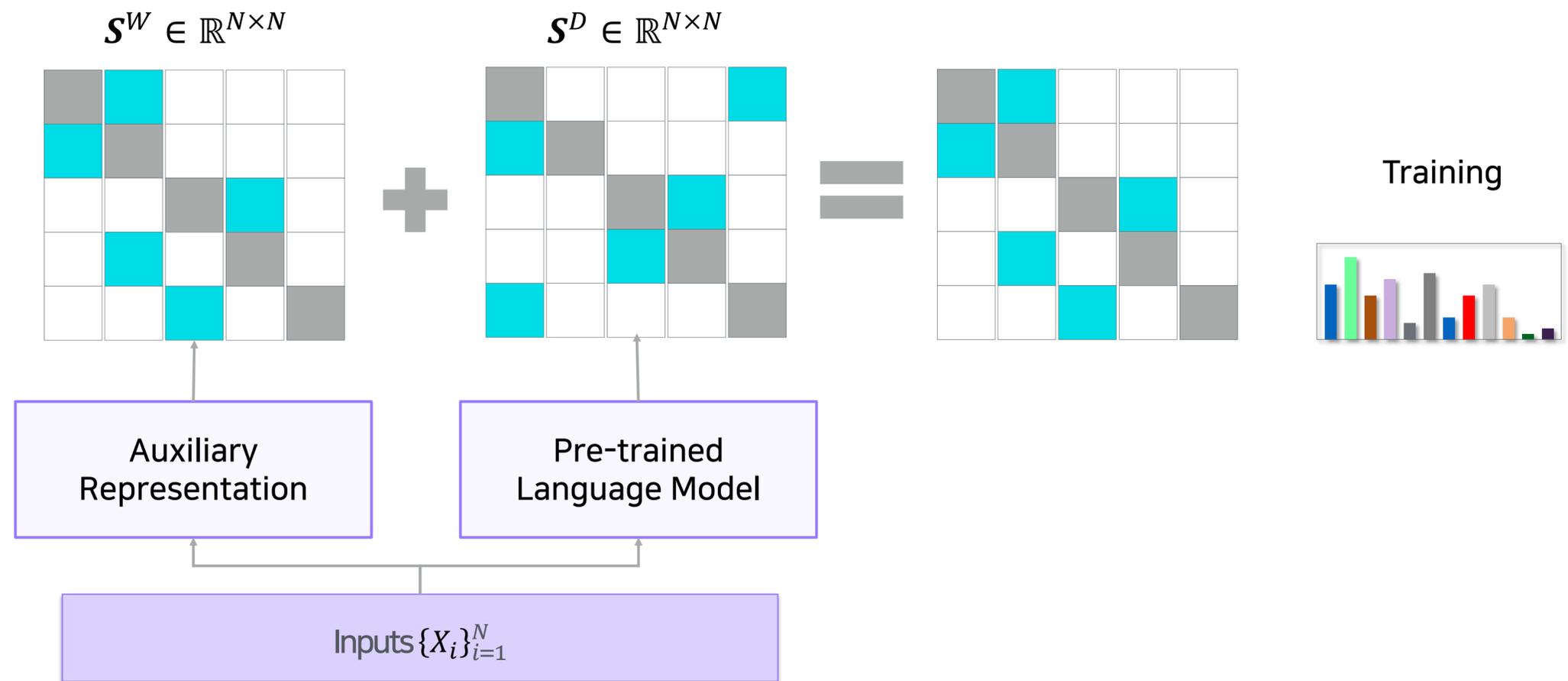
# 2.4 UNICON의 구조

Auxiliary Representation

Auxiliary Representation

초기에 불안정한 BERT를 가이드할 수 있는 보조적인 수단을 추가하자

Model	Accuracy
BM25	58.81%
TF-IDF	52.62%
GloVe	48.44
BOW	39.51



## 2.5 Benchmark Dataset 실험

### Experiments

- Intent Detection Datasets
- CLINC
- BANKING

Intent	Query
lost_luggage	I think my luggage has been lost
oil_change_when	How often should I change the oil
income	Do you know my work income

CLINC dataset samples

Datasets	# of domains	# of intents	Training	Validation	Test
CLINC	10	150	18,000	2,250	2,250
BANKING	1	77	9,003	1,000	3,080

Datasets distribution

# 2.5 Benchmark Dataset 실험

## Unsupervised Baselines

- DCN (Yang et al., ICML, 2017)
- SimCSE (Gao et al., EMNLP, 2021) NLP 에서 Contrastive Learning 을 사용한 최신 모델

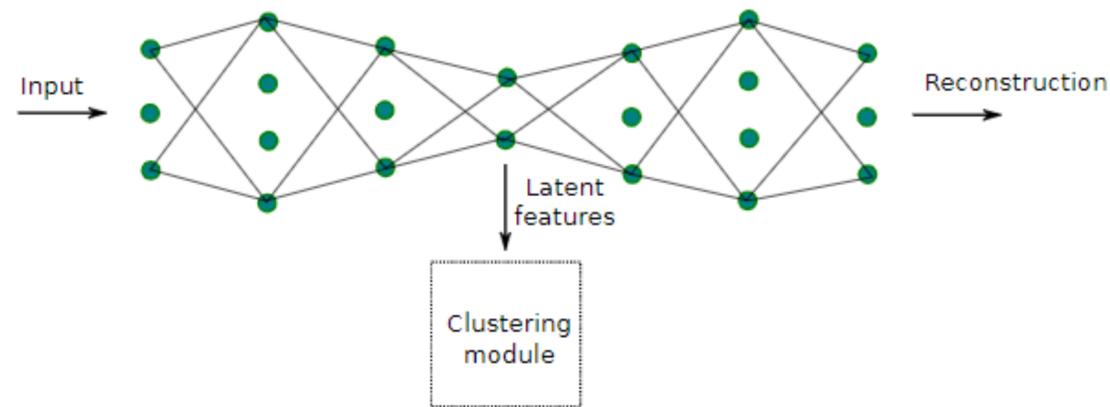
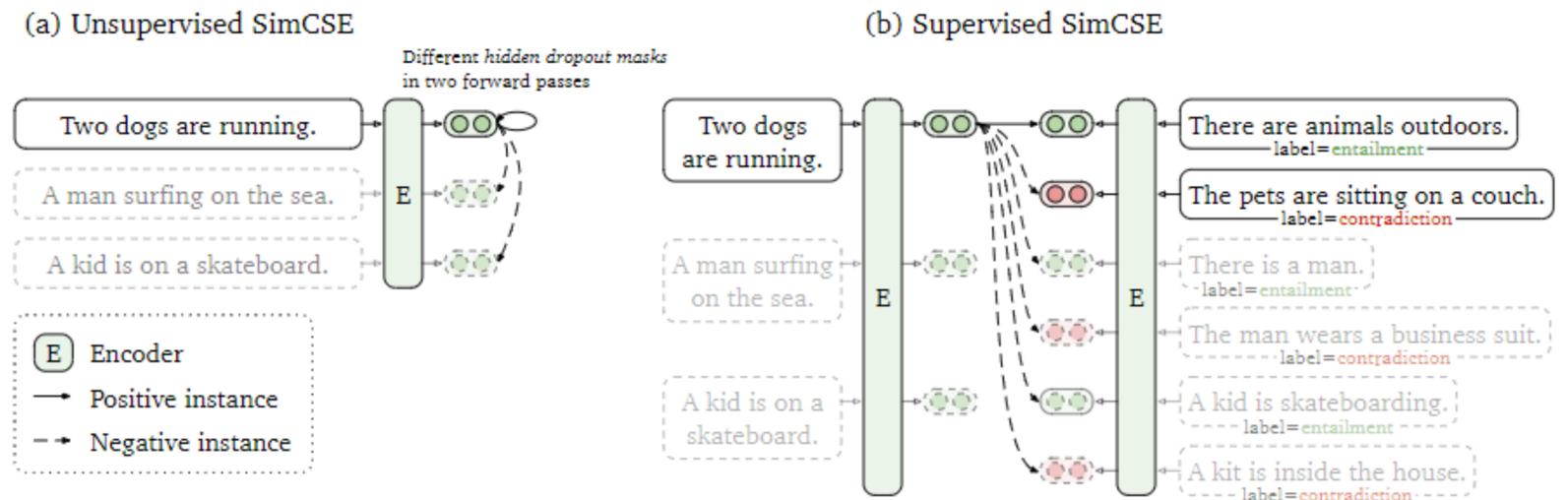


Figure 3. Proposed deep clustering network (DCN).

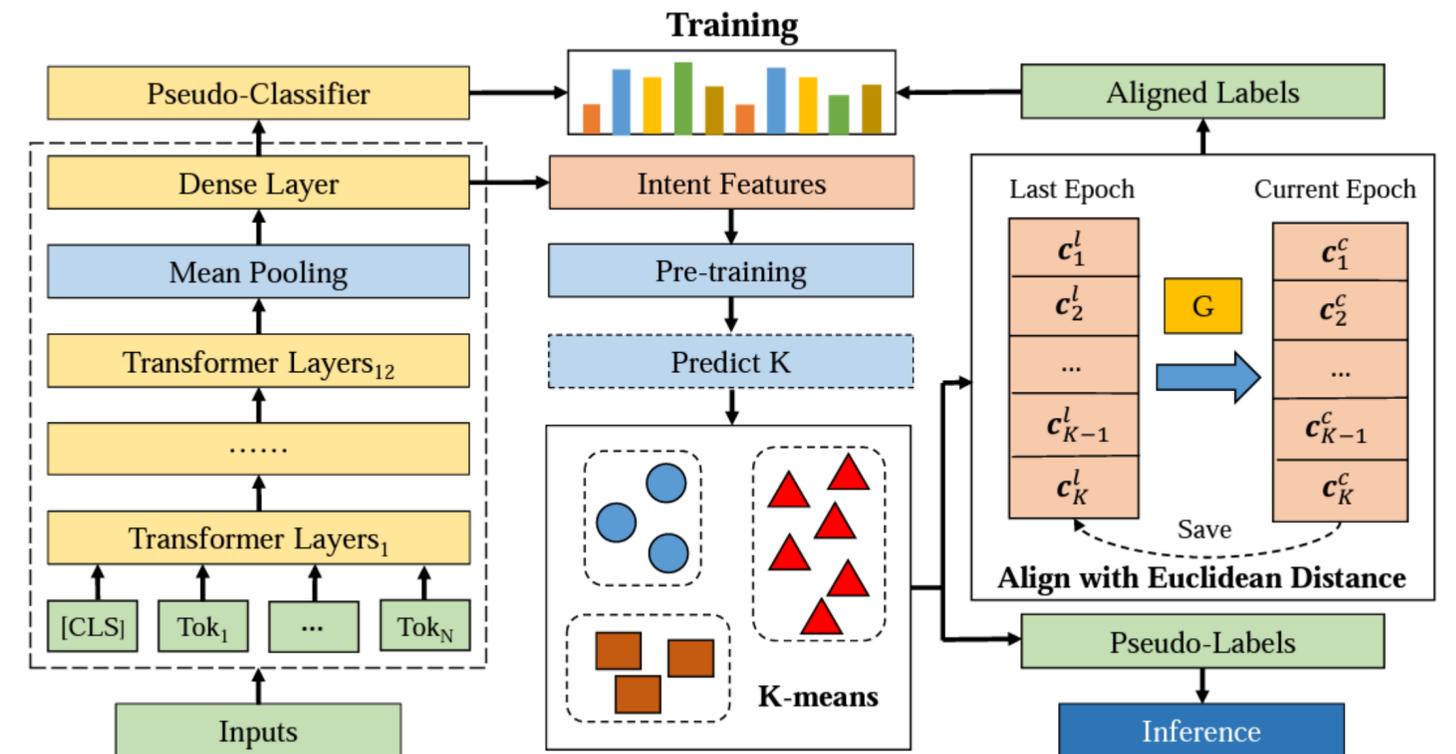
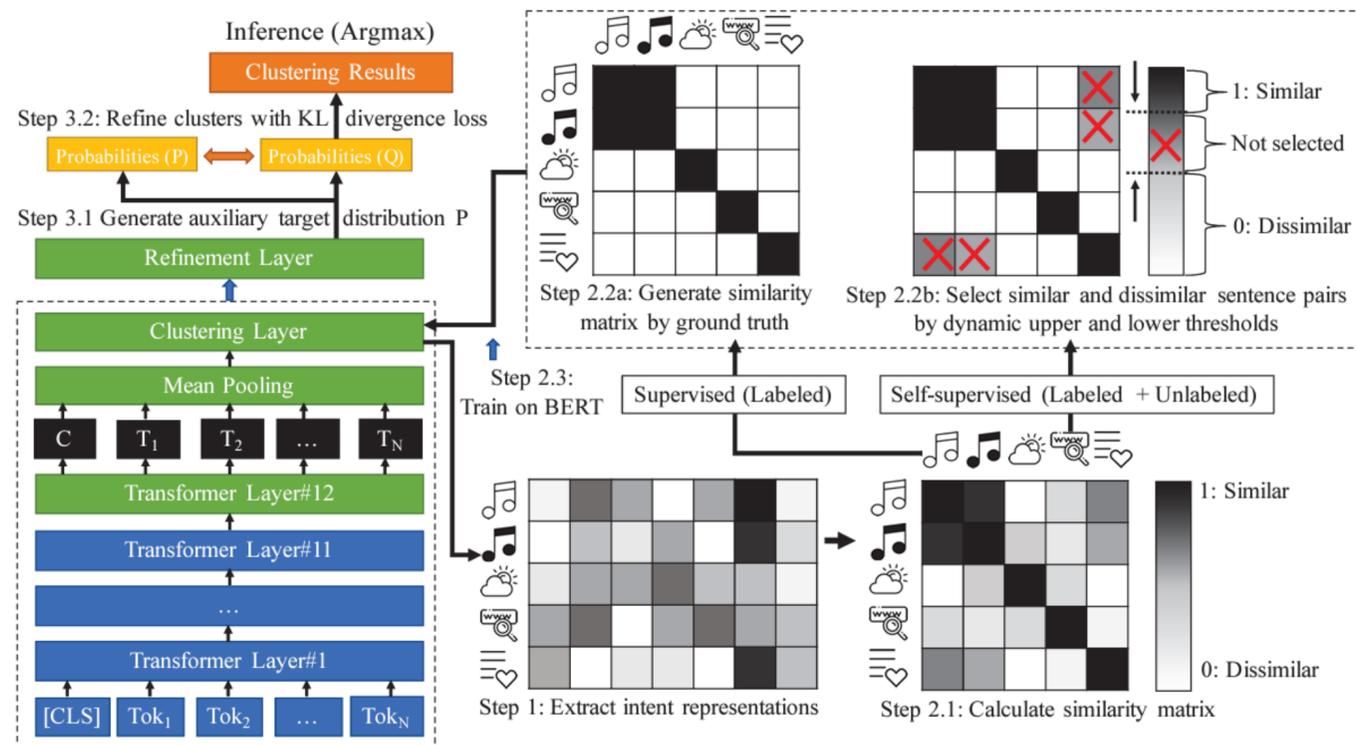


# 2.5 Benchmark Dataset 실험

## Semi-supervised Baselines

- CDAC+ (Lin et al., AAI 2020)
- DeepAligned (Zhang et al., AAI 2021)

Intent Discovery 에서 최신 모델



## 2.5 Benchmark Dataset 실험

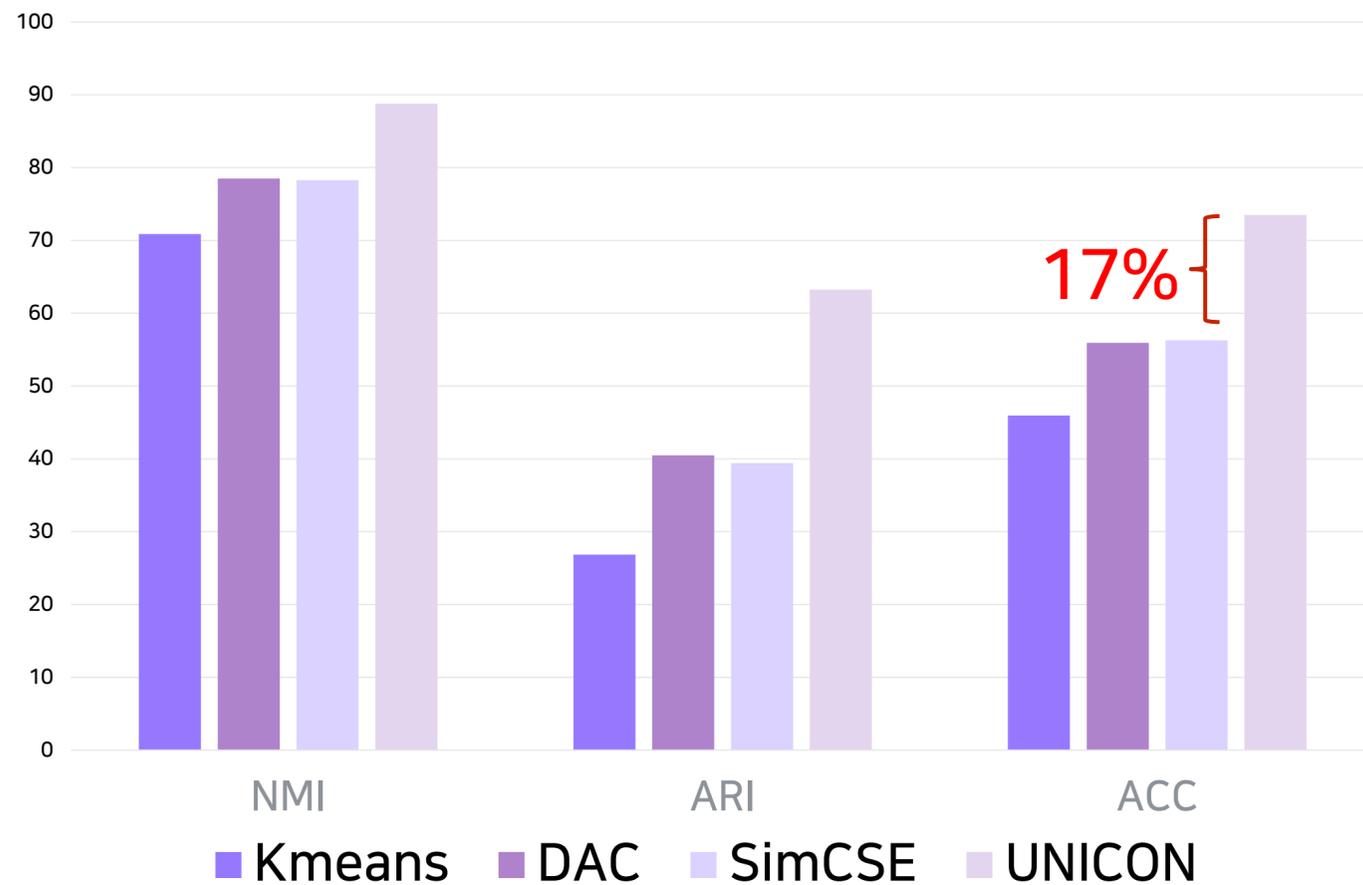
### Supervised Evaluation Metrics

- NMI (Normalized Mutual Information)
- ARI (Adjusted Rand Index)
- Accuracy (w/ hungarian alignment)

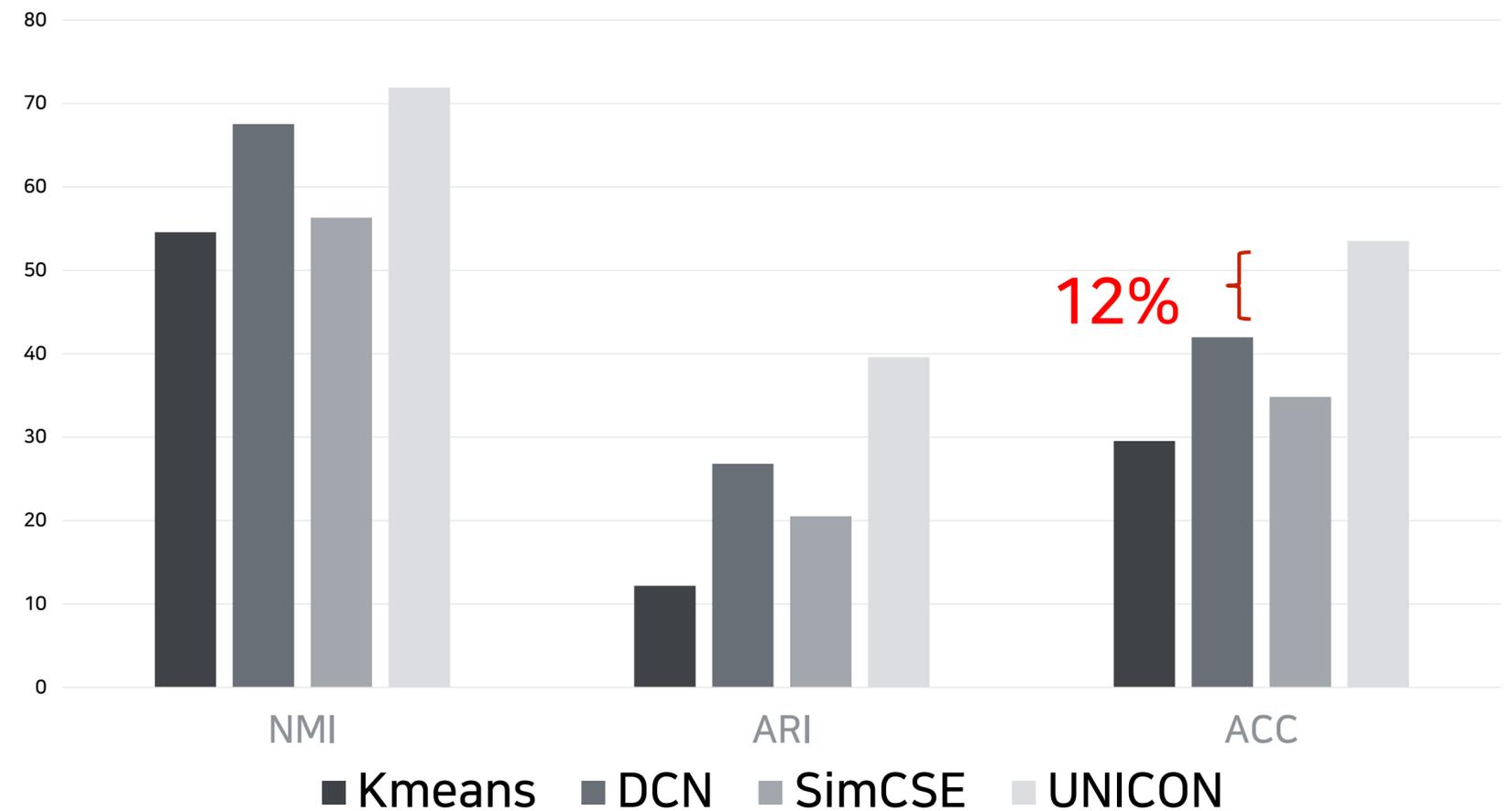
# 2.5 Benchmark Dataset 실험

Unsupervised:  
12-17% 성능 향상!

### CLINC Experiments



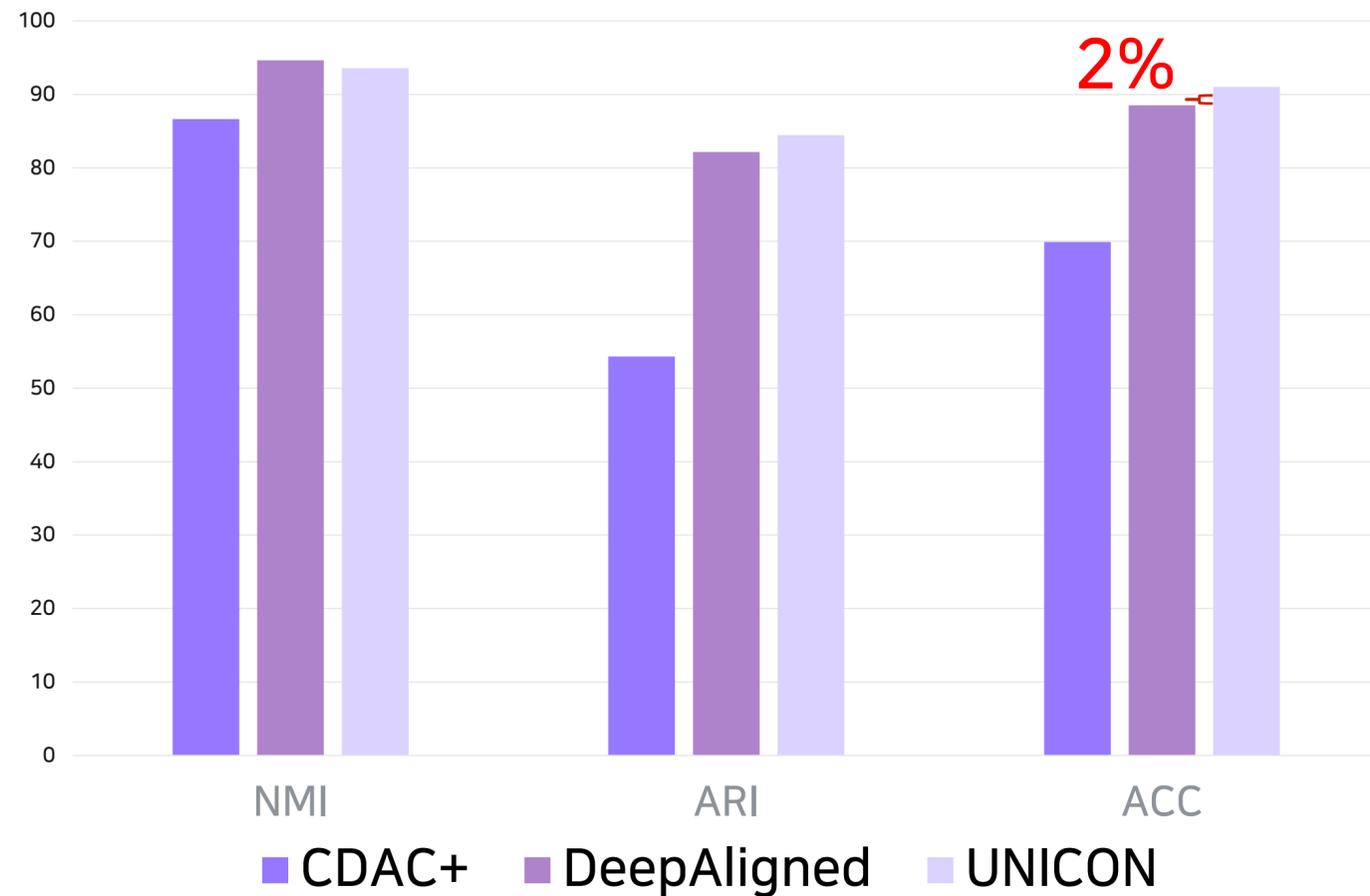
### BANKING Experiments



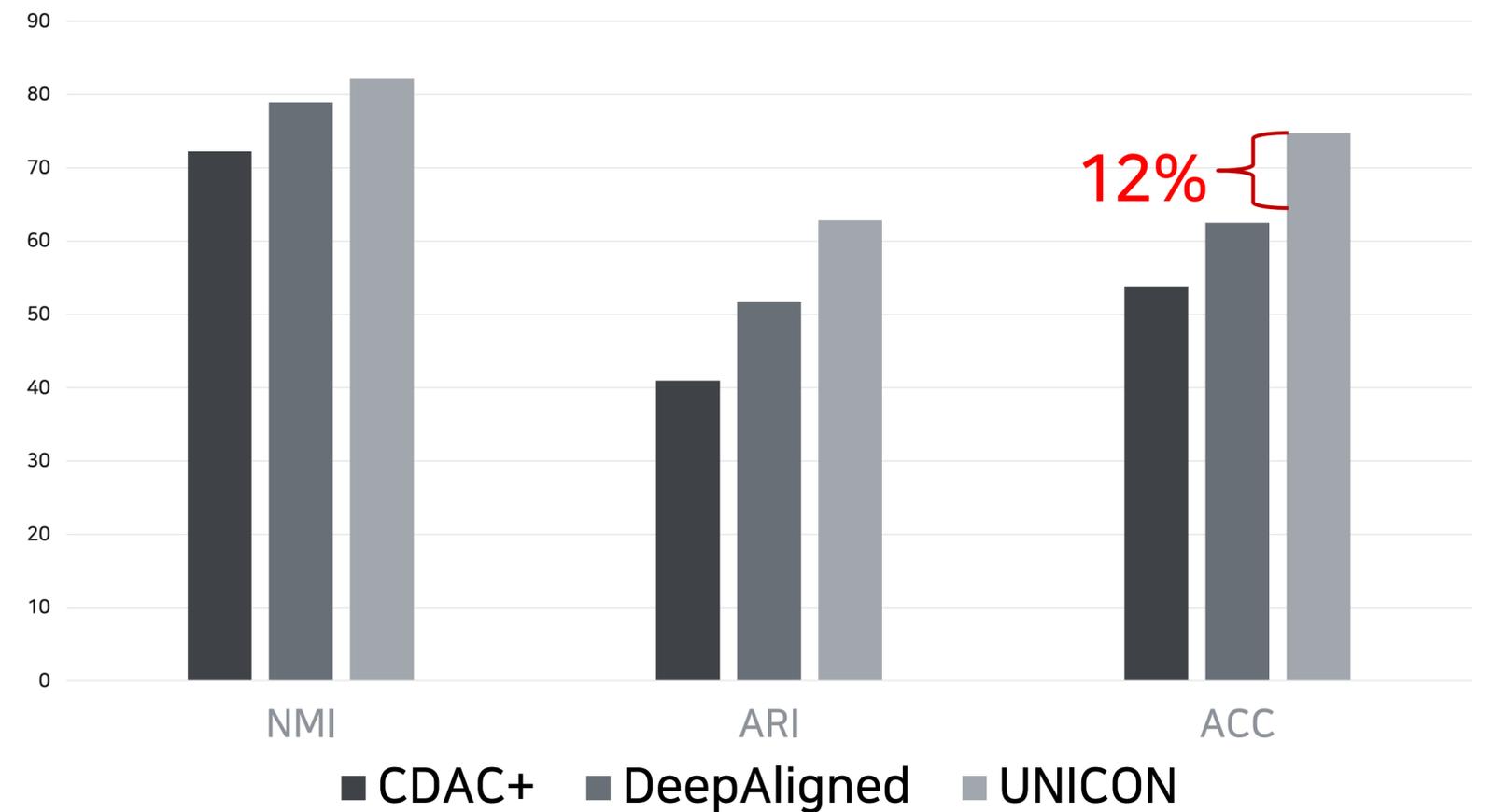
# 2.5 Benchmark Dataset 실험

Semi-supervised:  
2-12% 성능 향상!

### CLINC Experiments



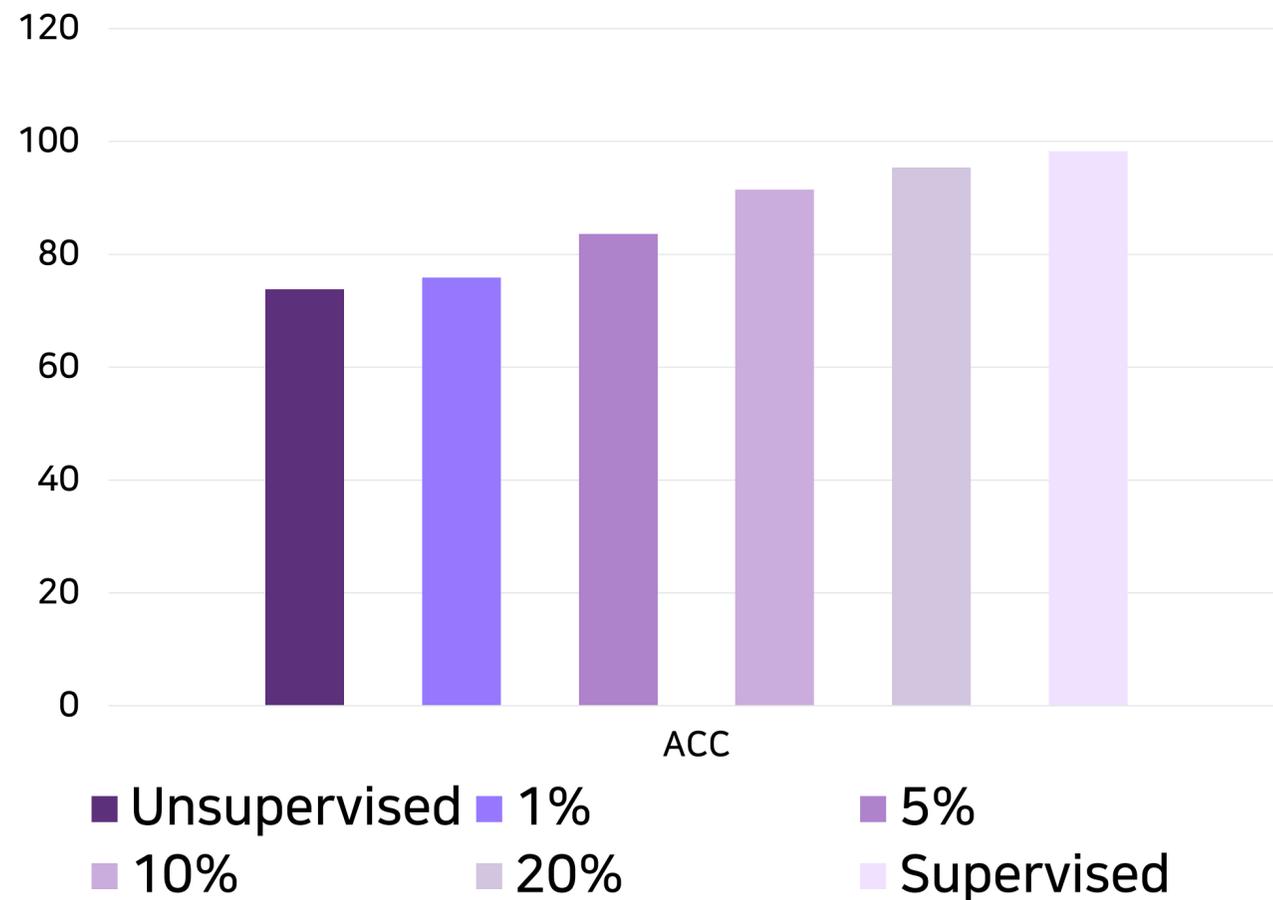
### BANKING Experiments



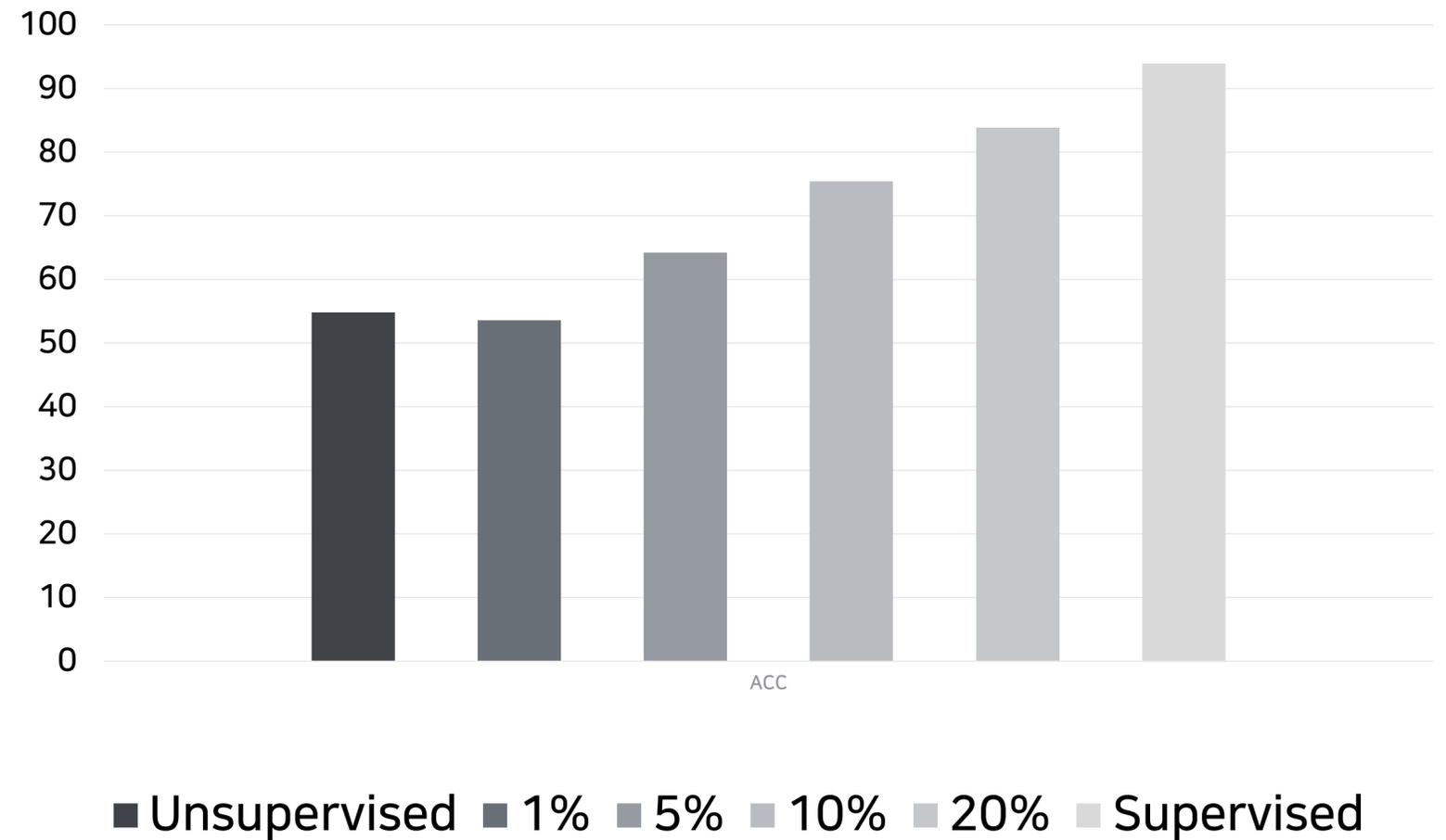
# 2.5 Benchmark Dataset 실험

- labeled data 를 추가할 수록 성능이 향상
- 20% labeled data 를 추가했을 때, supervised 대비 2%-10% 성능 차이

### CLINC Experiments

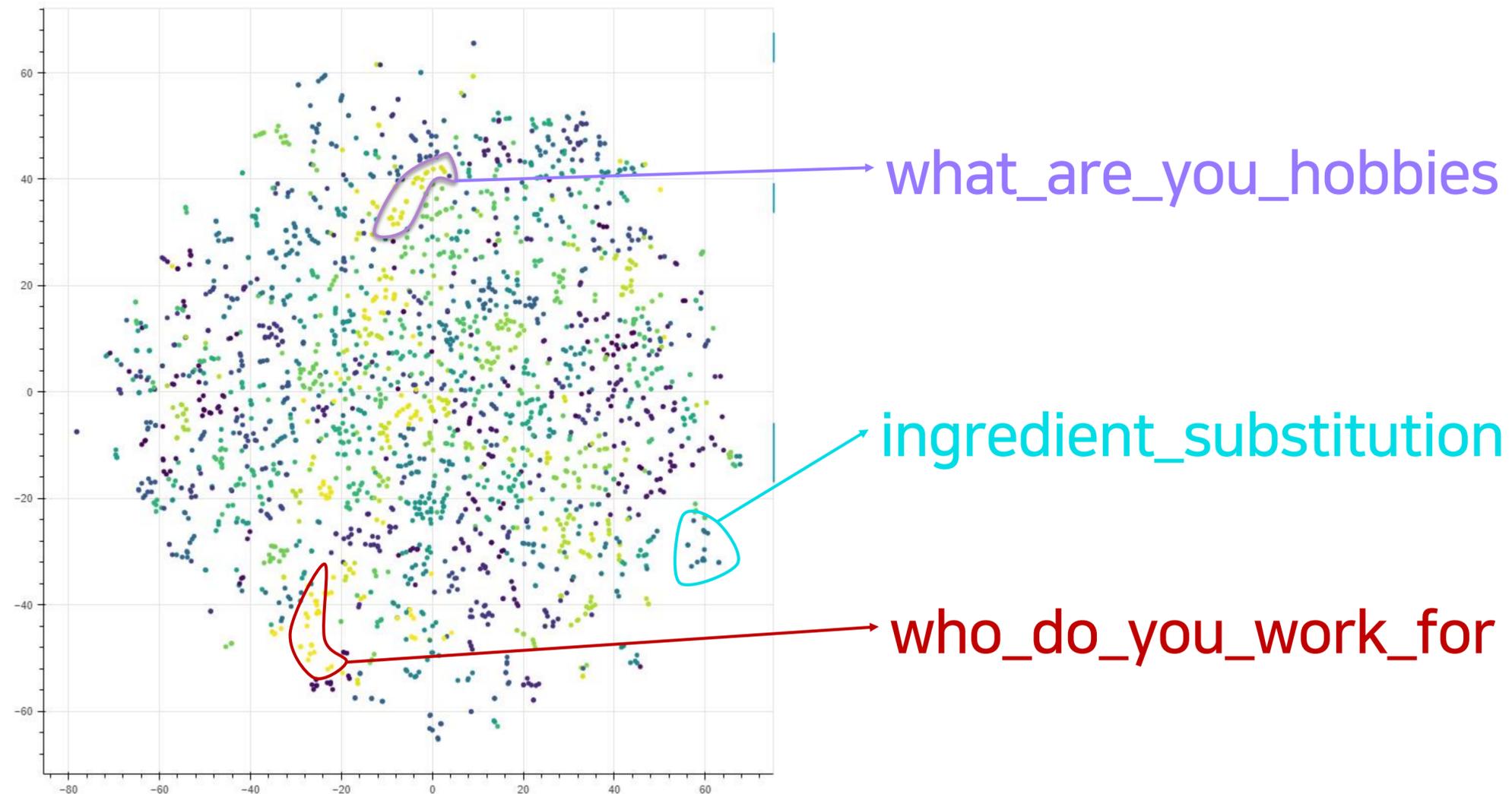


### BANKING Experiments



## 2.5 Benchmark Dataset 실험

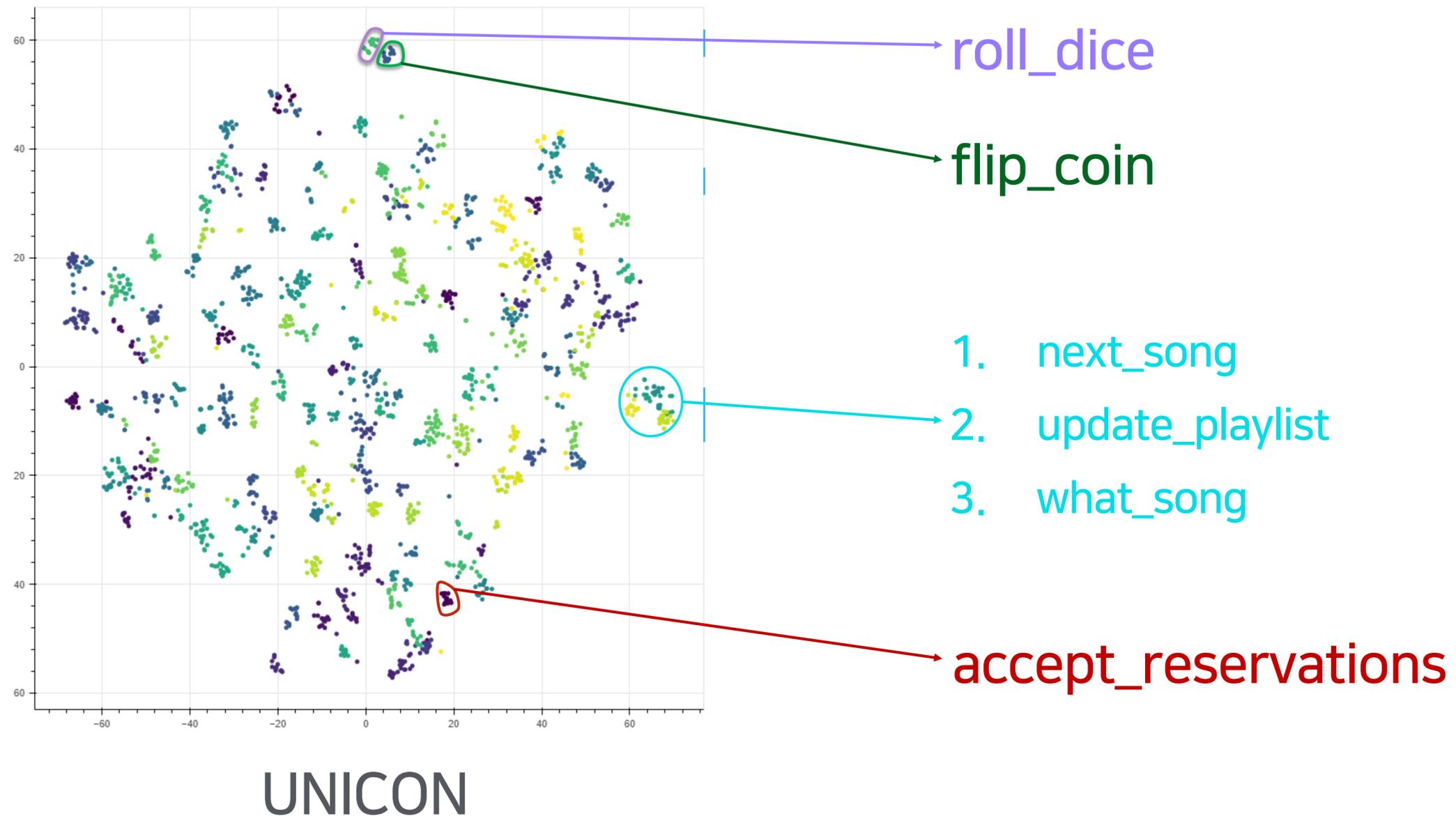
기존의 SimCSE (EMNLP 2021) 모델은 굉장히 퍼져있음



SimCSE (EMNLP 2021)

## 2.5 Benchmark Dataset 실험

UNICON의 경우 유사한 의도의 데이터끼리 뭉쳐있음



“**획기적으로** 성능을 향상시킨 **UNICON** 을 사용하면

데이터 설계 및 제작을 **자동화**할 수 있고

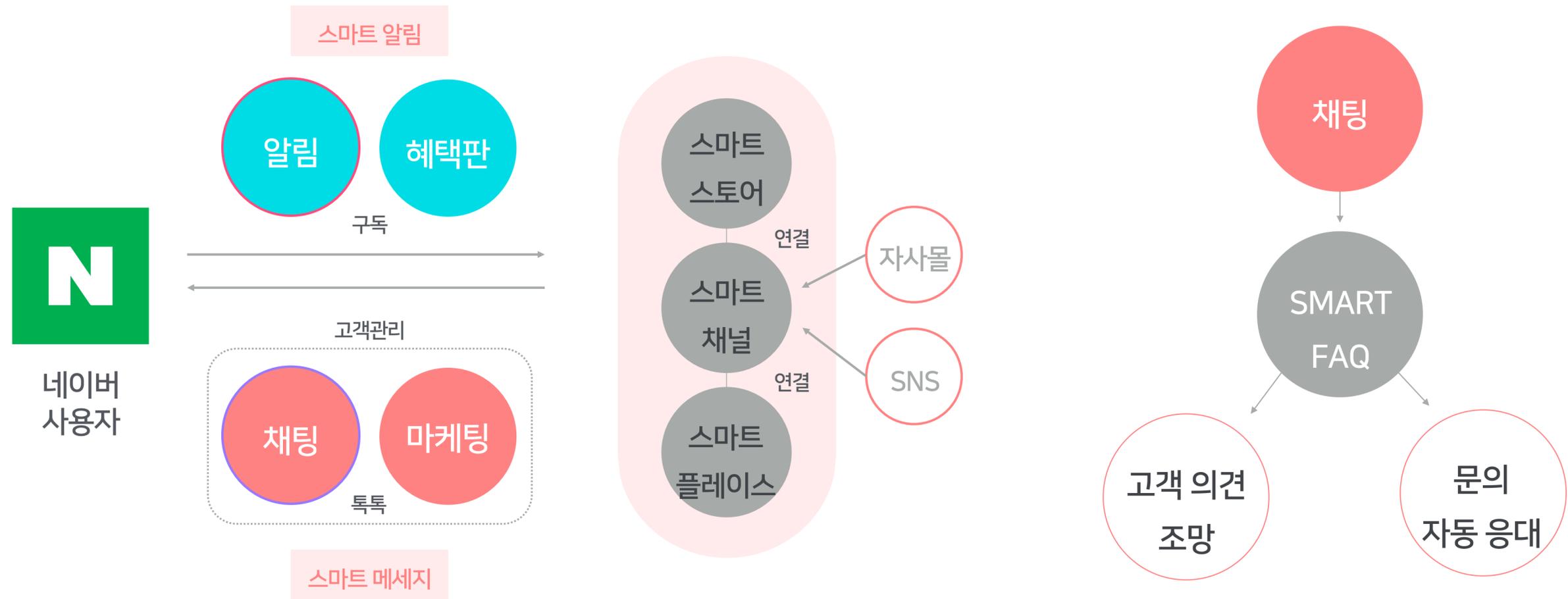
소모되는 **비용**을 **획기적으로** 줄일 수 있음”

# 3. 스마트스토어 톡톡 챗봇 자동응대 FAQ

# 3.1 스마트스토어 톡톡 챗봇 자동응대 FAQ

AI 지원 기능으로 고객 문의를 효율적으로 관리

서비스 타겟: 네이버 스마트 스토어 SME 판매원들



# 3.1 스마트스토어 톡톡 챗봇 자동응대 FAQ

## 네이버 톡톡

11개 항목

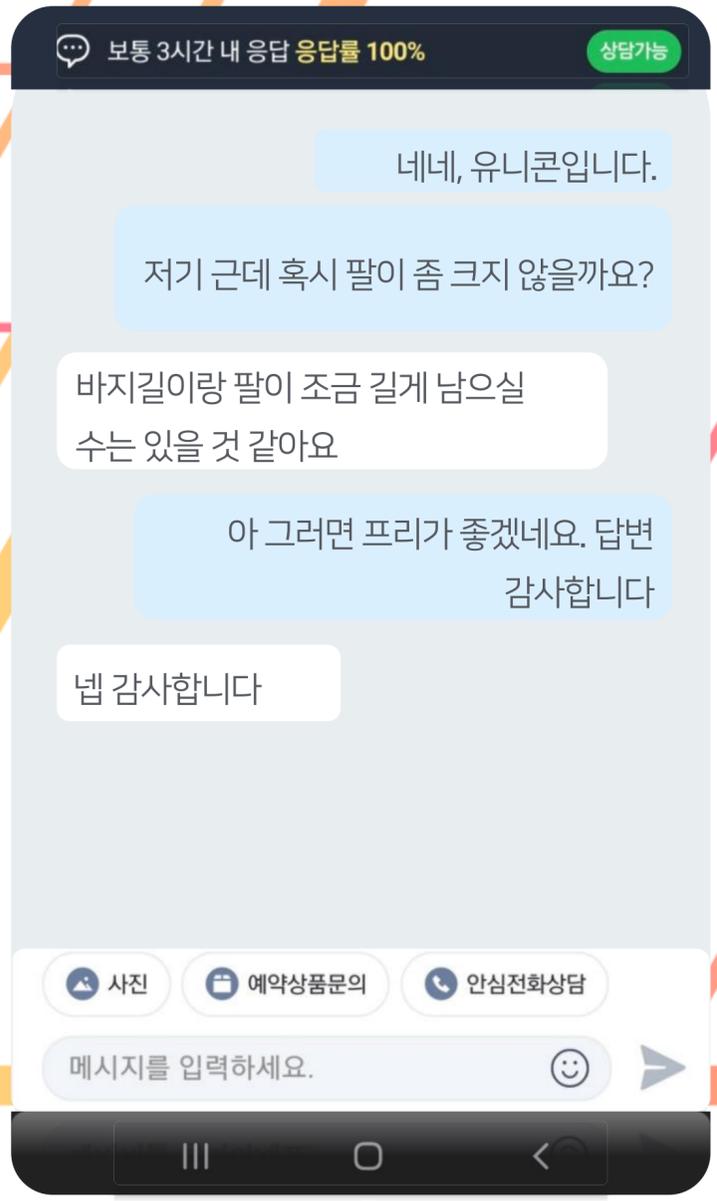
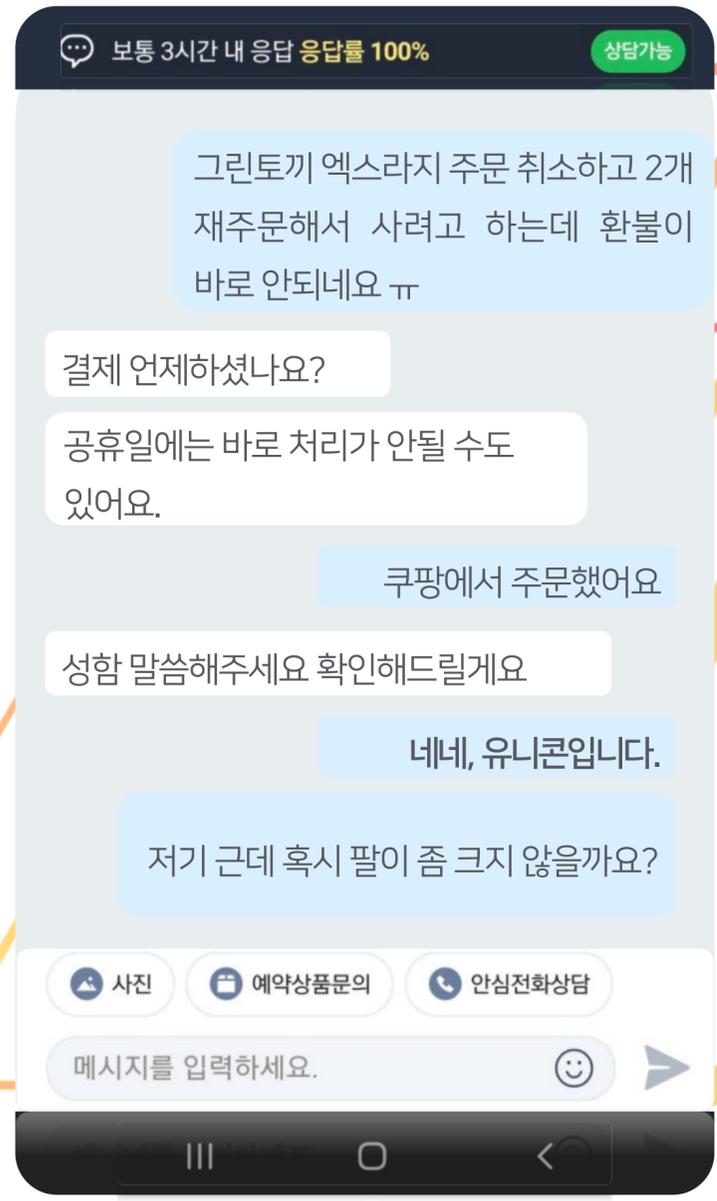
패션, 패션잡화, 육아, 가구, 전자기기, 식료품, 스포츠, 생활, 레저, 도서, 화장품

익명화 진행

익명화, 비식별화 진행

위 대화는 가상의 시나리오로 구성된 대화이며, 모델에 사용된 톡톡 데이터셋은 "일반적인 상품 문의 등" 에 관한 것으로, 사적 대화로 구성되지 않았습니다.

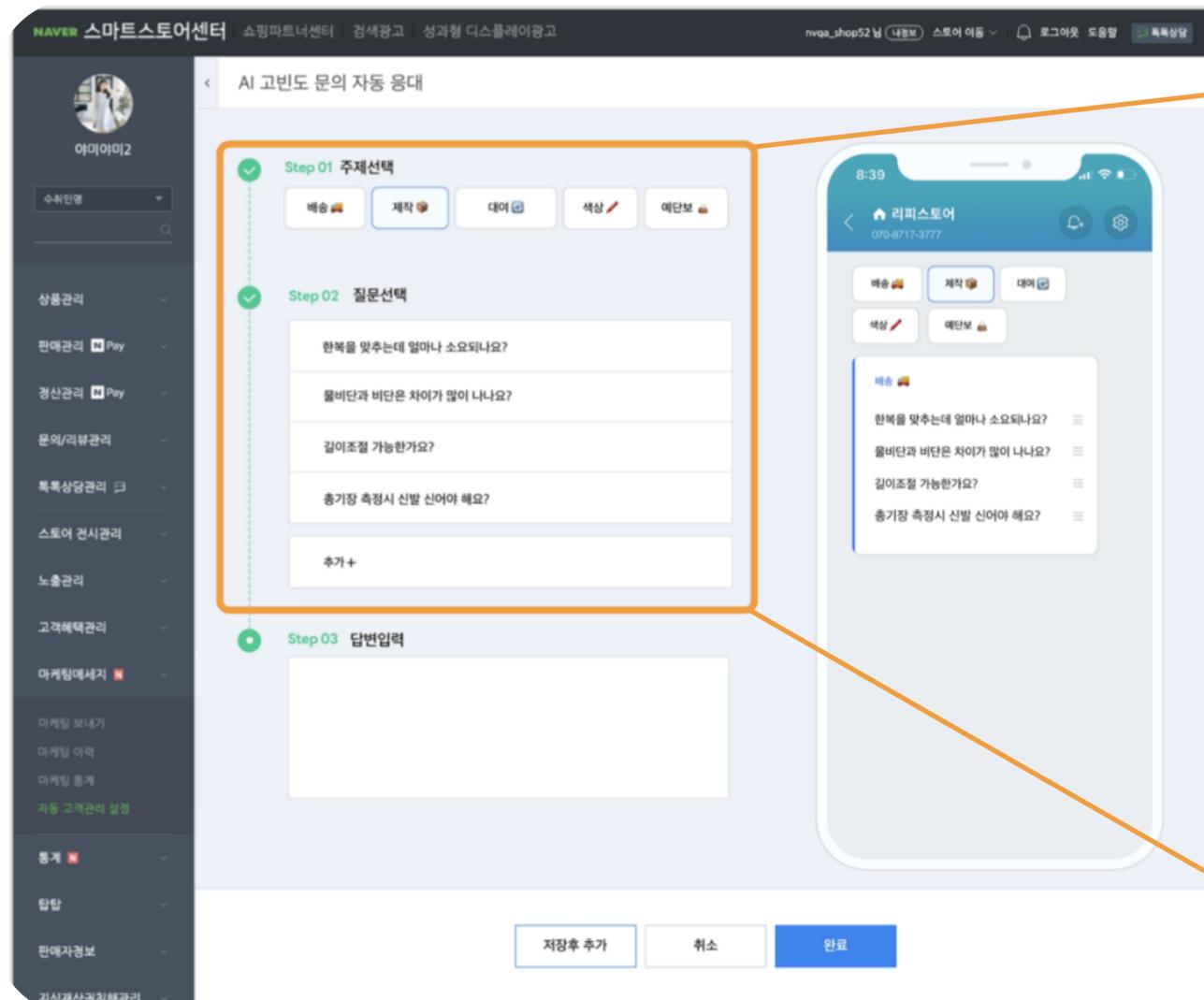
### 네이버 톡톡 대화 예제



# 3.1 스마트스토어 톡톡 챗봇 자동응대 FAQ

## 서비스 예제

고객들이 궁금해하는 점 요약



### Step 1 주제선택

배송    제작    대여    색상    예단보

### Step 2 질문선택

한복을 맞추는데 얼마나 소요되나요?

물비단과 비단은 차이가 많이 나아요?

길이 조절 가능한가요?

총기장 측정시 신발 신어야 해요?

# 3.1 스마트스토어 톡톡 챗봇 자동응대 FAQ

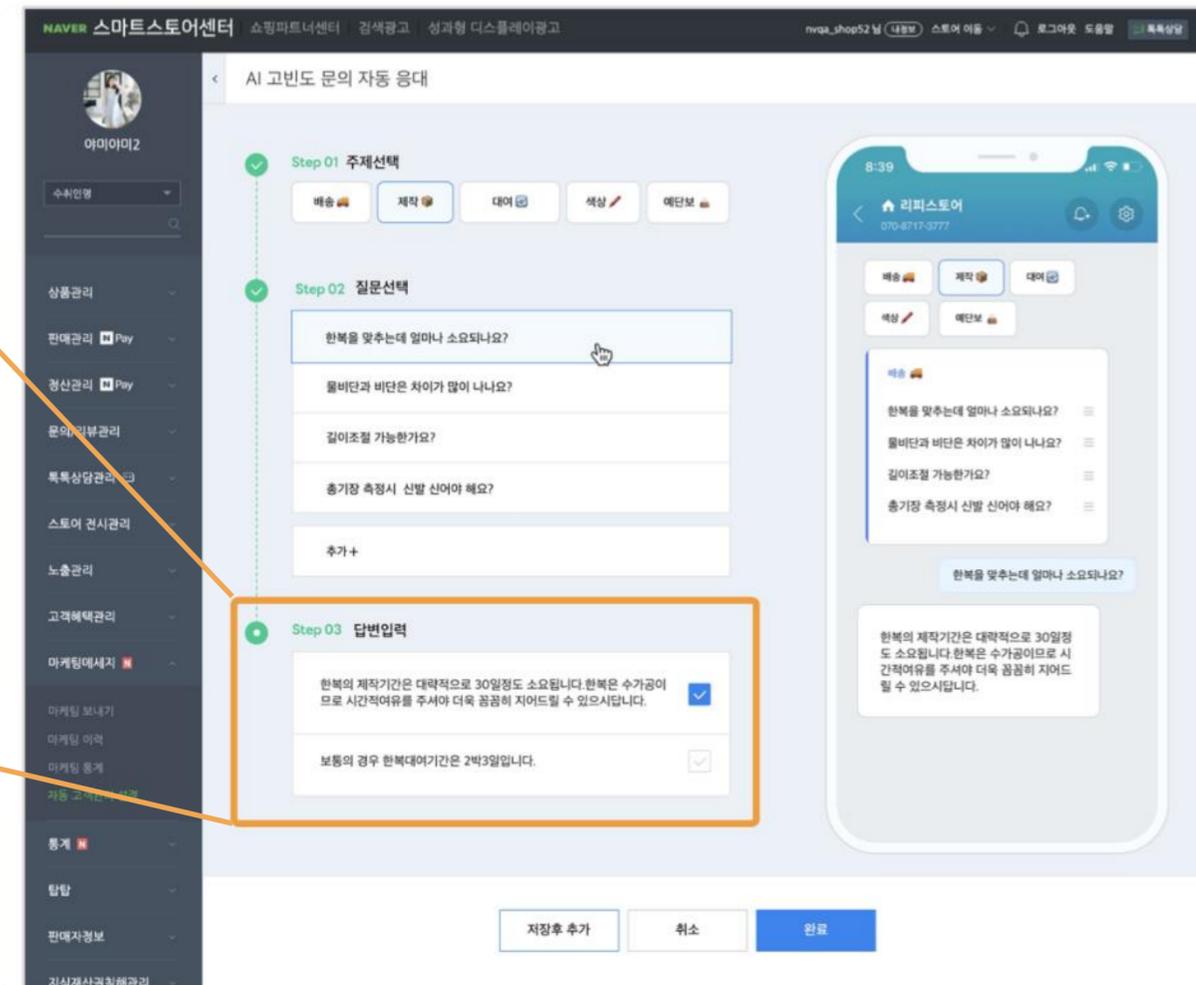
## 서비스 예제

판매원이 미리 등록한 답변으로 자동 응대

### Step 3 답변입력

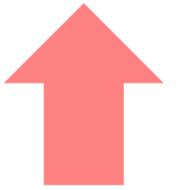
한복의 제작기간은 대략적으로 30일 정도 소요됩니다.  
한복은 수가공이므로 시간적 여유를 주셔야 더 꼼꼼히  
지어드릴 수 있습니다.

보통의 경우 한복 대여기간은 2박 3일입니다.

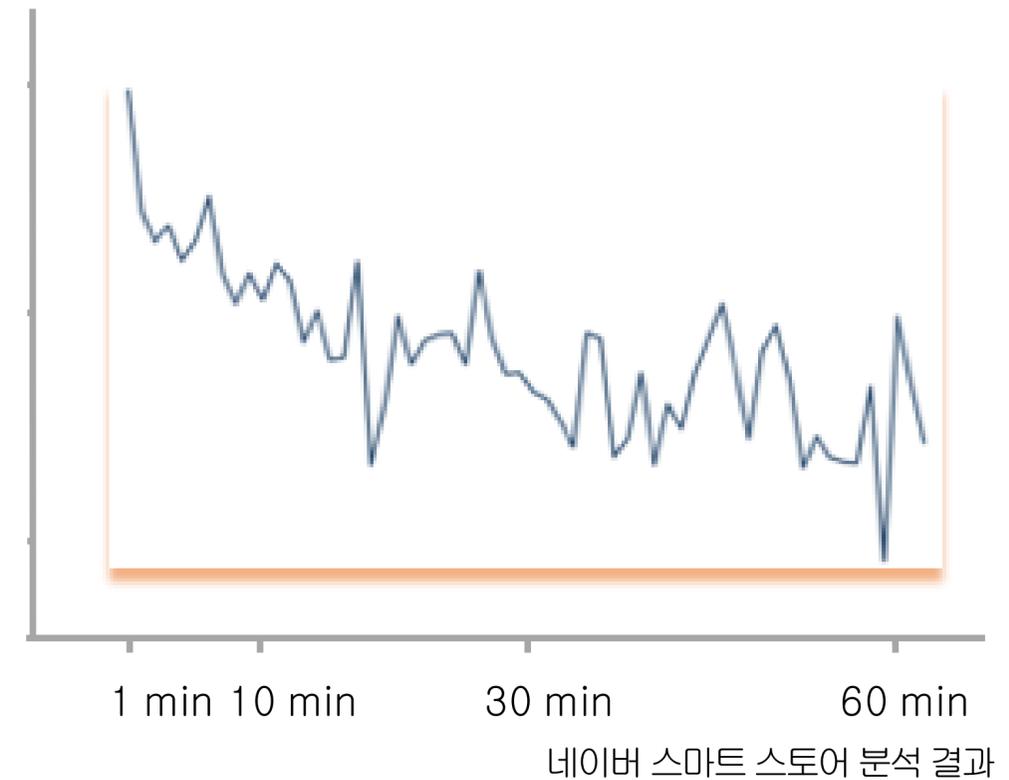


# 3.1 스마트스토어 톡톡 챗봇 자동응대 FAQ

고객문의에 빠르게 대답할 수록 매출로 이어질 확률이 크다



구매 전환율

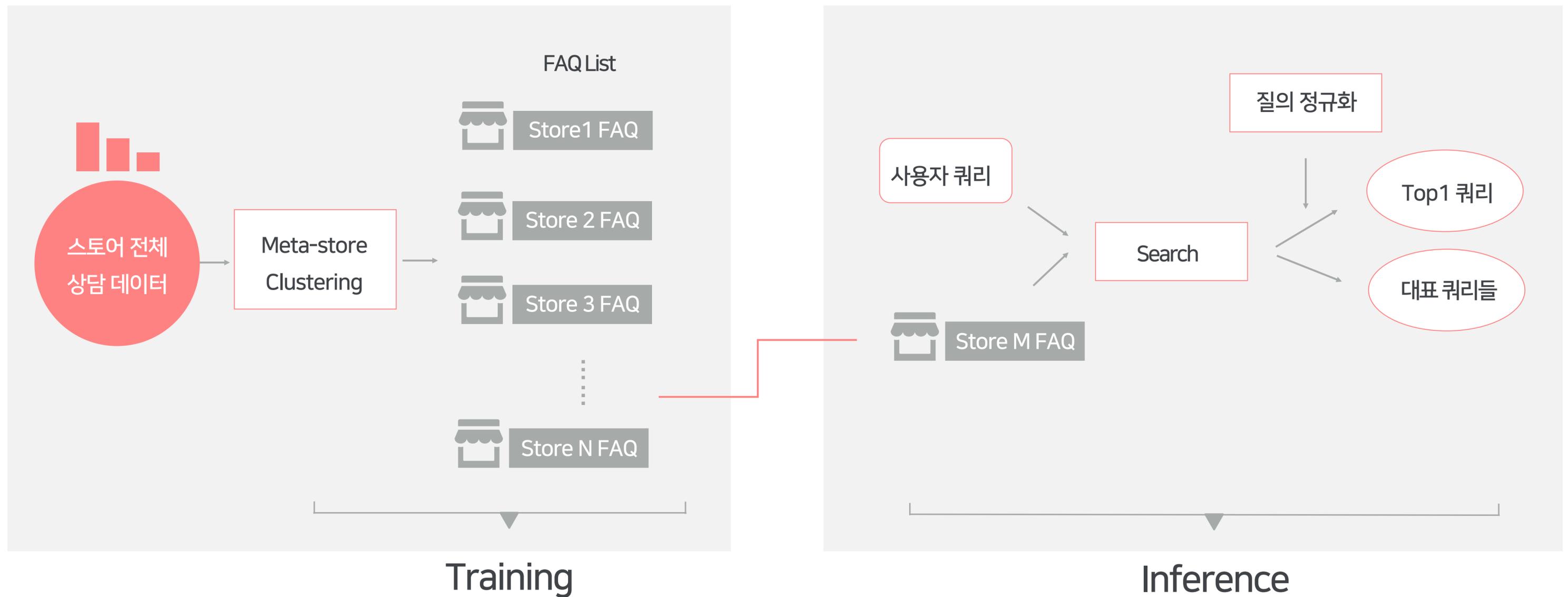


응답 시간

“스토어 맞춤형 응대 서비스를 어떻게  
기술적으로 구현할 수 있을까?”

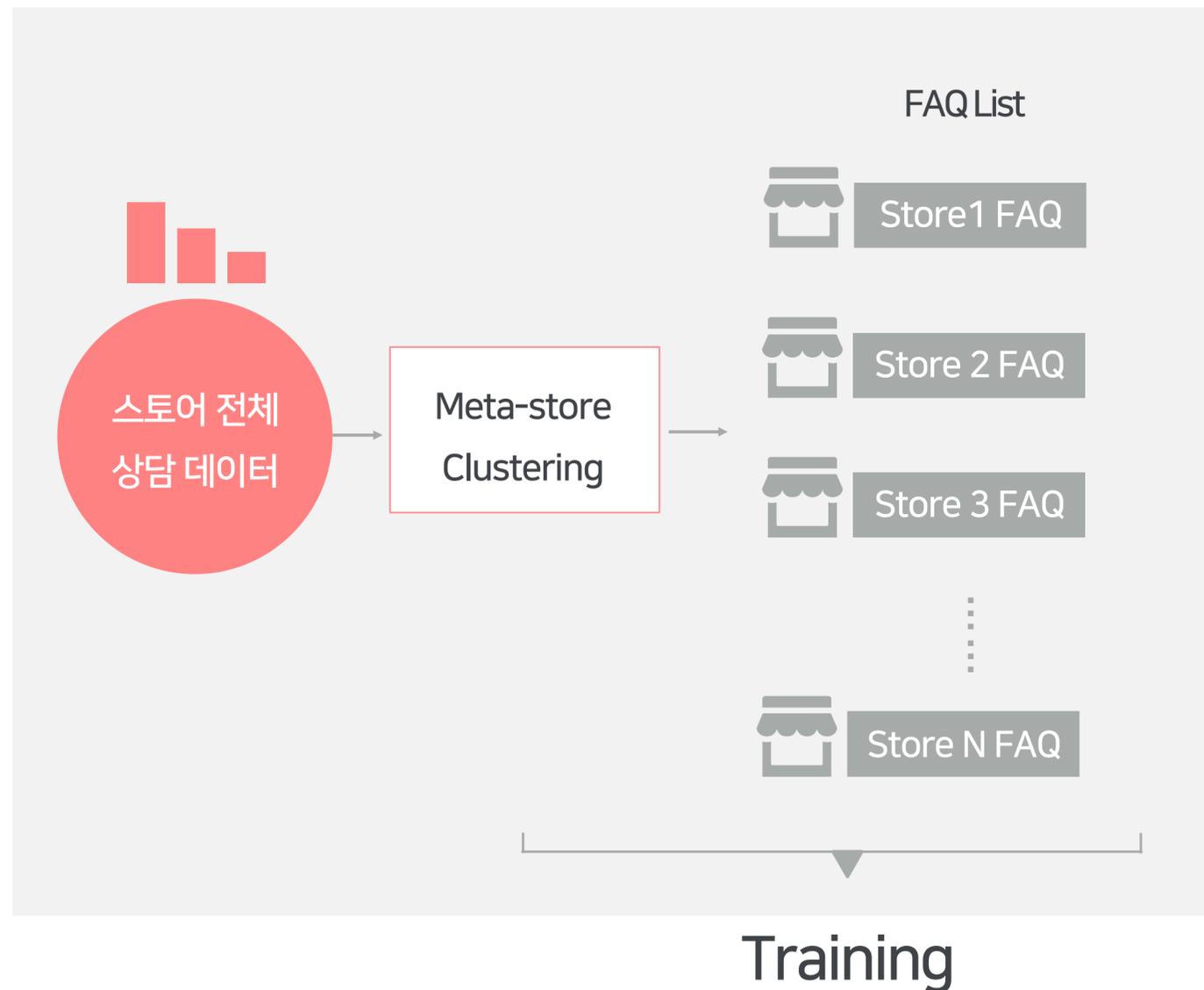
# 3.1 스마트스토어 톡톡 챗봇 자동응대 FAQ

## Overall Framework



# 3.2 자동응대-FAQ: Training

## Mate-store Clustering



Meta-store clustering

- 각 스토어 별로 모델을 학습 → 개별 모델 퀄리티 저하
- 전체 데이터로 학습 (카테고리 별)

구현 디테일

- 사용자 쿼리가 model의 input: 질의 이전의 대화 history를 이어 붙임
- 형태소 기반의 tokenizer를 sparse representation 구할 때 활용

# 3.2 자동응대-FAQ: Training

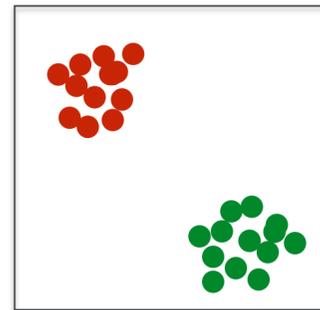
## Silhouette Score

Reference label이 필요 없는 intrinsic metric 인 실루엣 스코어 활용

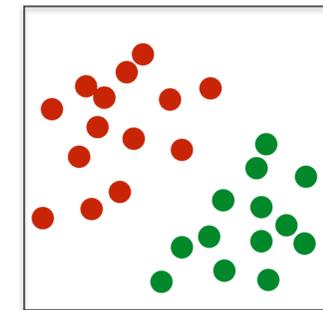
$$a(i) = \frac{1}{|C_i|-1} \sum_{j \in C_i, i \neq j} d(i, j)$$

$$b(i) = \min_{k \neq i} \frac{1}{|C_k|} \sum_{j \in C_k} d(i, j)$$

$$s(i) = \frac{b(i) - a(i)}{\max\{a(i), b(i)\}}$$



score ↑



score ↓

➡ 또한 클러스터링 이후 coherent하지 않은 cluster들을 제거 하는데 사용

talktalk-19 classification dataset

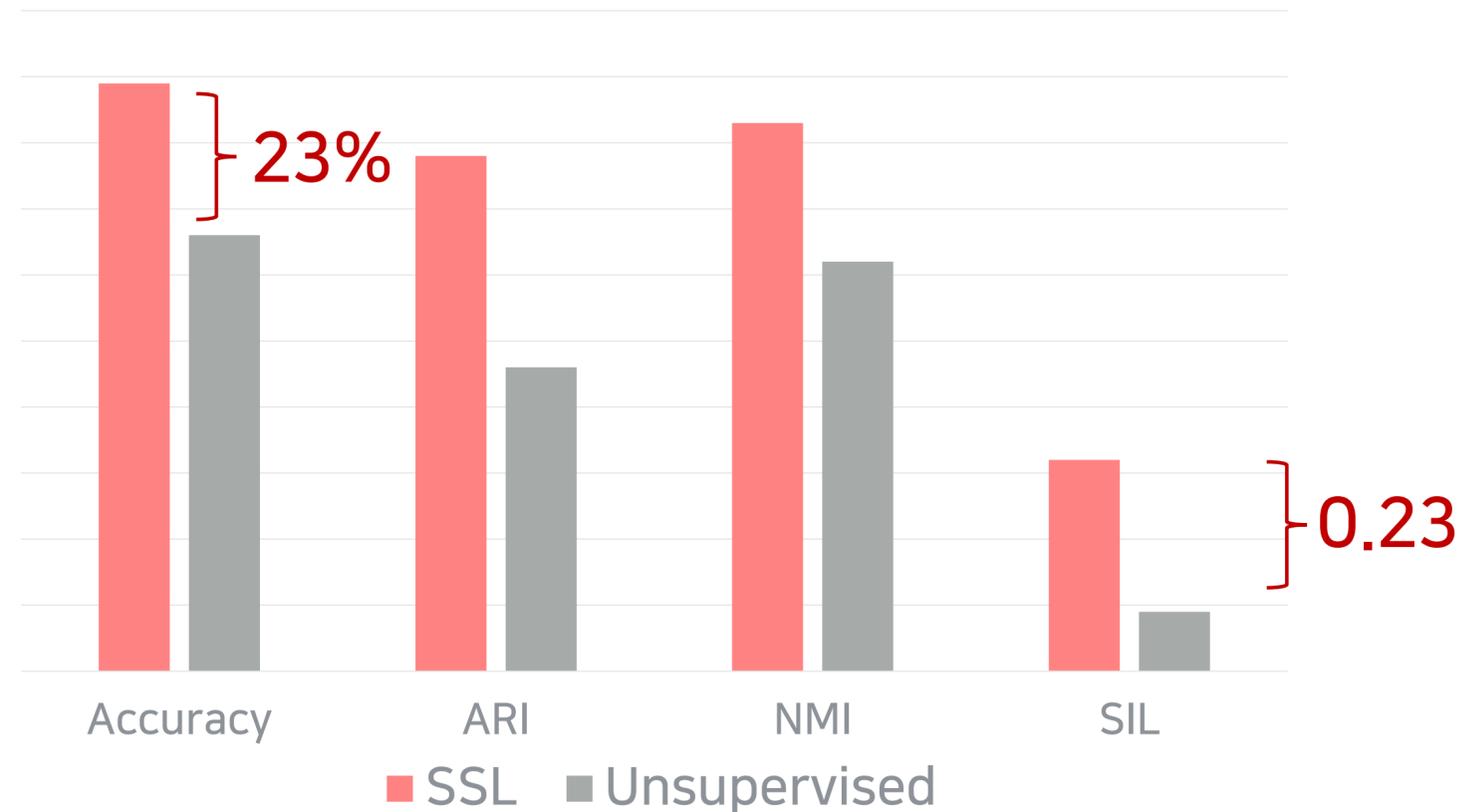
- 19개 label 체계 (e.g., 배송문의, 상품문의) 를 갖고 있으며 학습용 40k, 평가용 4k로 구성
- 평가용 데이터를 reference label이 필요한 metric을 측정할 때 활용 (Accuracy, NMI, ARI)

**“어떻게 하면 모델 성능을 더 끌어올릴 수 있을까?”**

## 3.2 자동응대-FAQ: Training

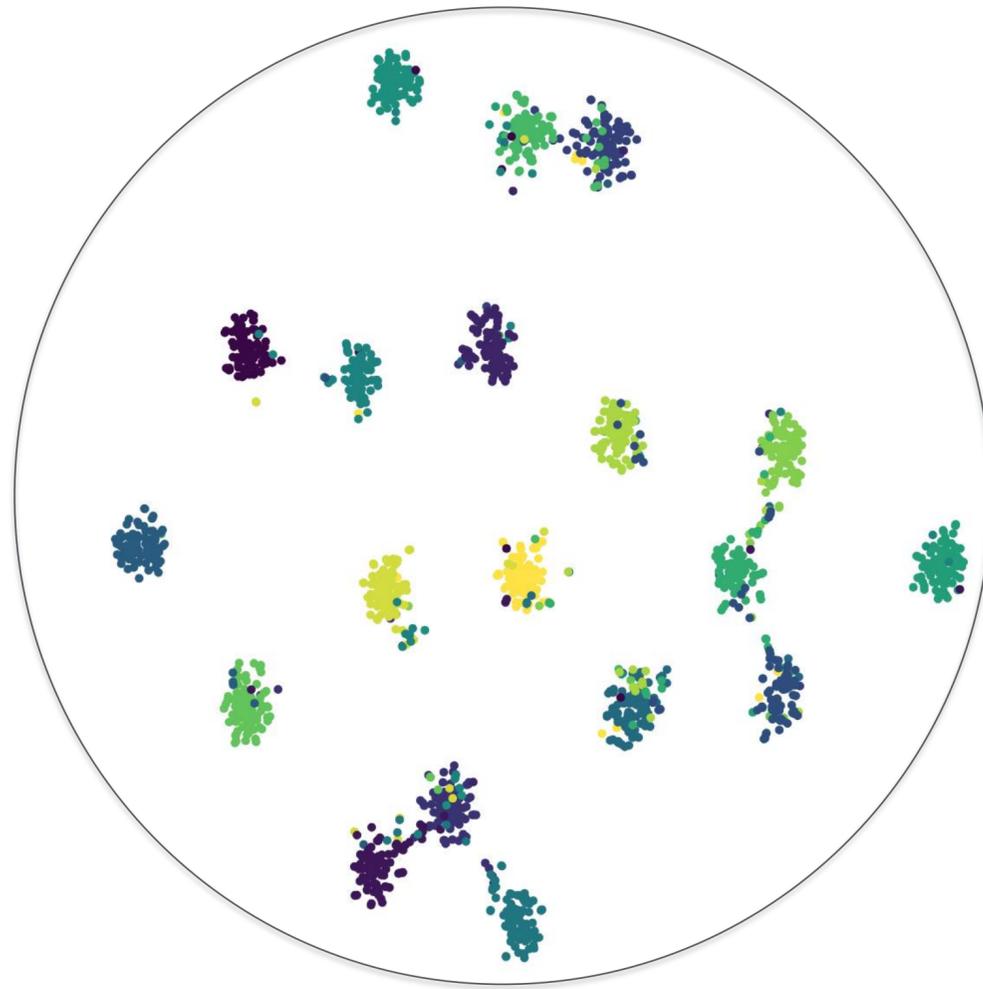
### Semi-Supervised Learning

- Talktalk19 학습셋을 활용, SSL 진행
- 각 batch마다 labeled sample이 균등하게 분포
- 도메인이 너무 다르지 않다는 가정하의 어느 분류 데이터도 활용 가능



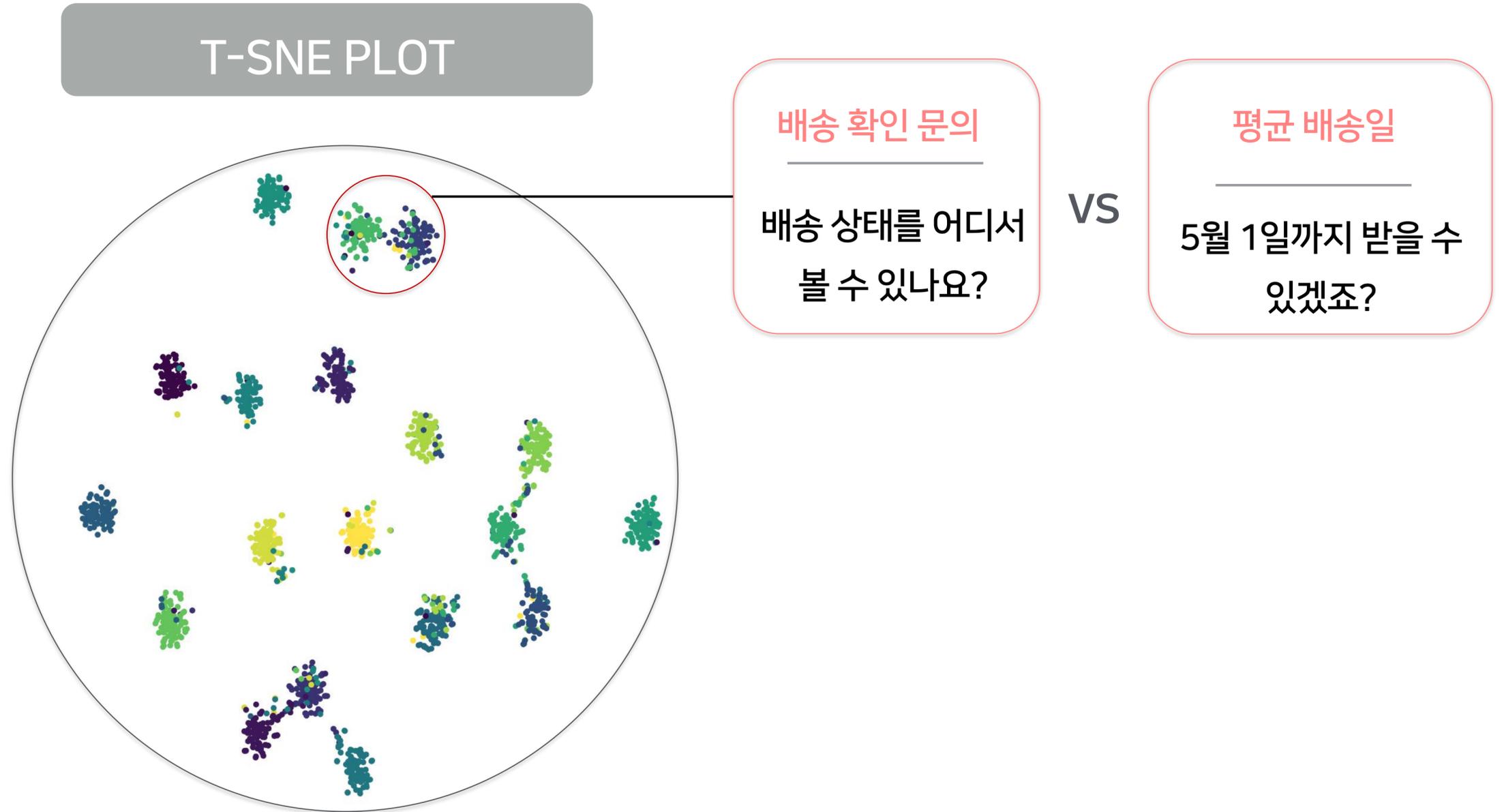
## 3.2 자동응대-FAQ: Training

T-SNE PLOT



Talktalk19 test set에 대한 T-SNE plot  
(label마다 색깔 할당)

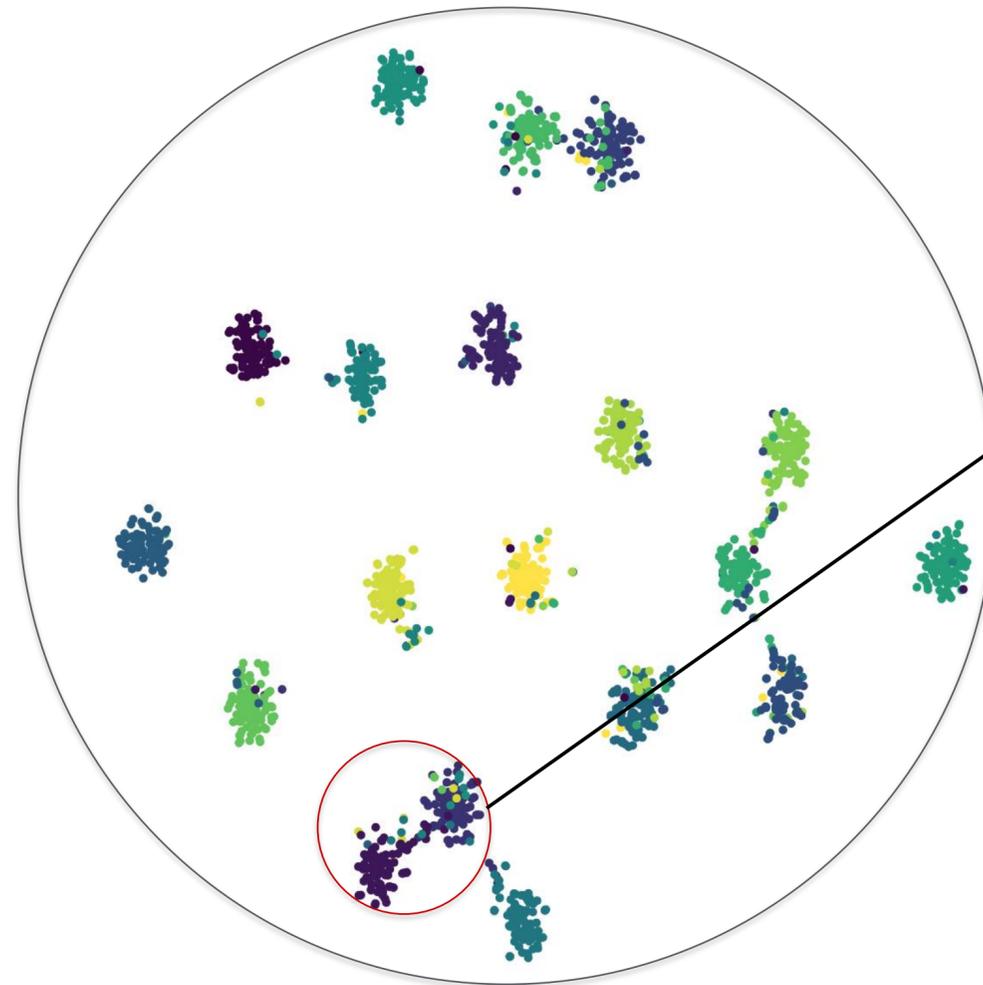
# 3.2 자동응대-FAQ: Training



Talktalk19 test set에 대한 T-SNE plot  
(label마다 색깔 할당)

# 3.2 자동응대-FAQ: Training

T-SNE PLOT



상품 정보 문의  
양말 사이즈 문의 좀  
드릴게요

VS

상품 하자 문의  
양말 크기가 서로  
다르네요

Talktalk19 test set에 대한 T-SNE plot  
(label마다 색깔 할당)

## 3.2 자동응대-FAQ: Training

클러스터링 결과

### 사이즈 문의

History:

User: 블랙 스몰사이즈 가겠습니다.

Seller: 스몰이면 28사이즈 인가요?

User query: 30입니다

⋮

History:

User: 바지 스몰사이즈 48키로 괜찮을까요?

Seller: 티업뎃되었어요 30400이구 스몰하심

딱맞으실것같아용

User query: 미뭉은 크겠쥬??

## 3.2 자동응대-FAQ: Training

클러스터링 결과

### 배송 문의

History:

User: 방금 결제했는데 혹시 이번주까지  
받을수 있을까요??ㅠㅠ

Seller: 위 상품은 이번주는 어려울듯 하구요

User query: 네에..

⋮

History:

User: 그럼 지연 없으니 이번주에 받아볼수  
있나요?

Seller: 네네 이번주에 가능하실 것 같아요!

User query: 감사합니다!!!

## 3.2 자동응대-FAQ: Training

클러스터링 결과

### 관세 문의

History:

User: 관세포함 아니었나요? πππ

Seller: 독일에서 발송하면 관세 혜택되어  
위 금액으로 가능하세요

User query: 아아 ㅎㅎ

⋮

History:

User: 관세가 추가발생하게 되나요

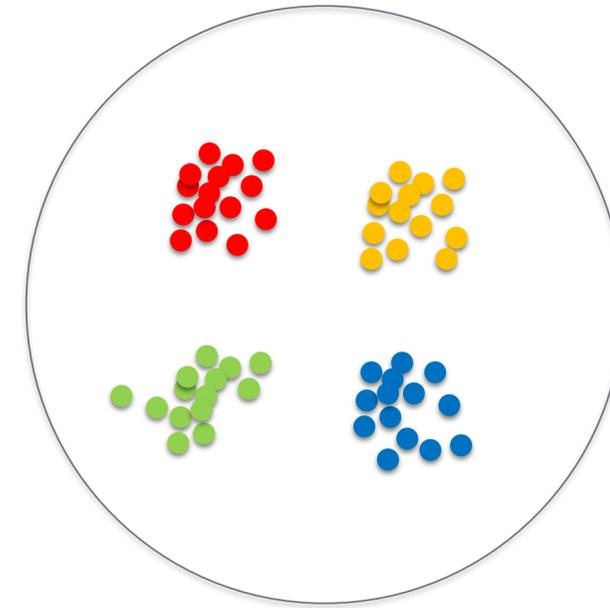
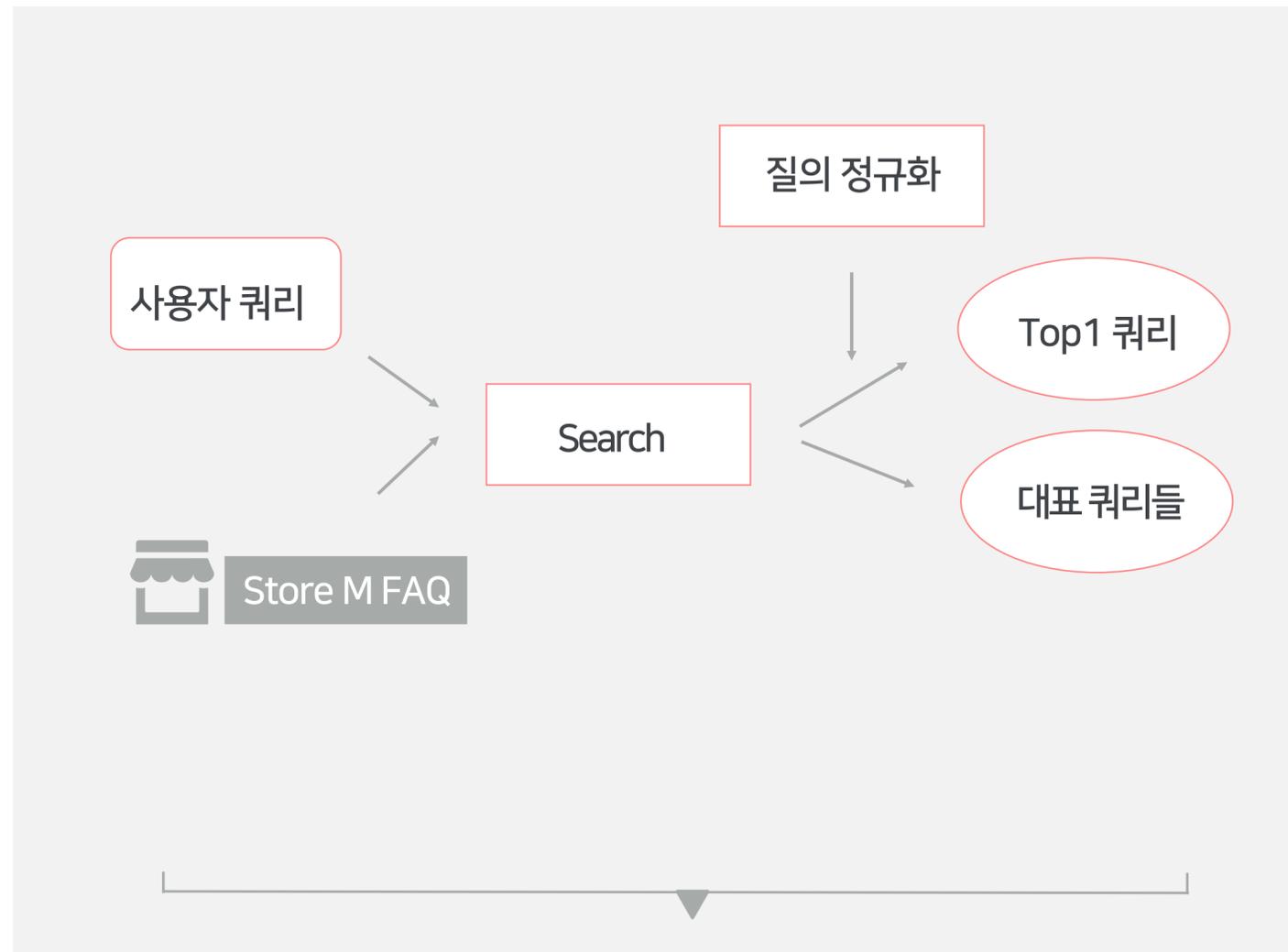
Seller: 네 영국 발송으로 세금이  
발생합니다π

User query: 세금이 얼마인가요?

“이제 어떻게 서비스로 구현하지?”

# 3.3 자동응대-FAQ: Inference

## Similarity Search

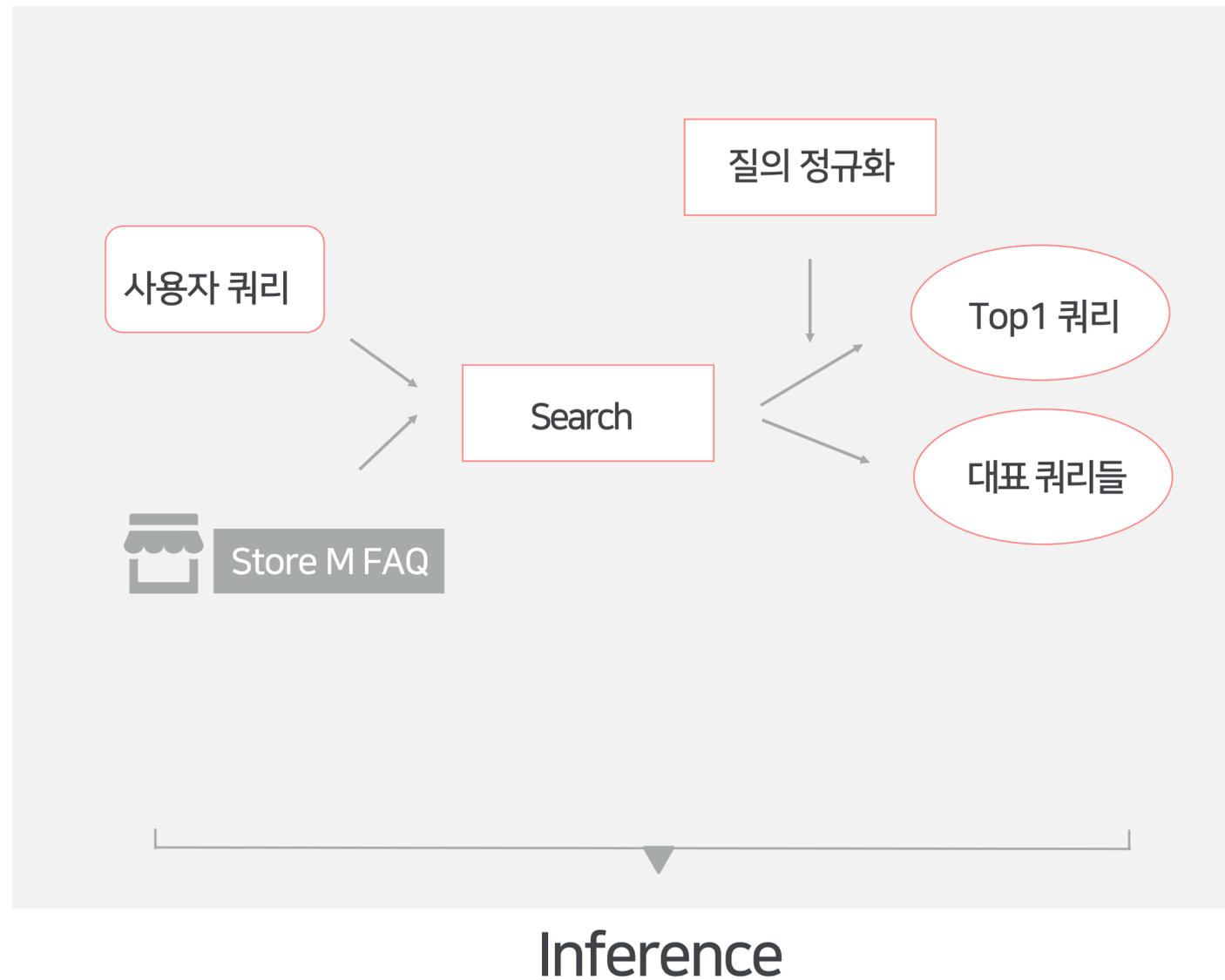


Meta-Store Cluster Data

Inference

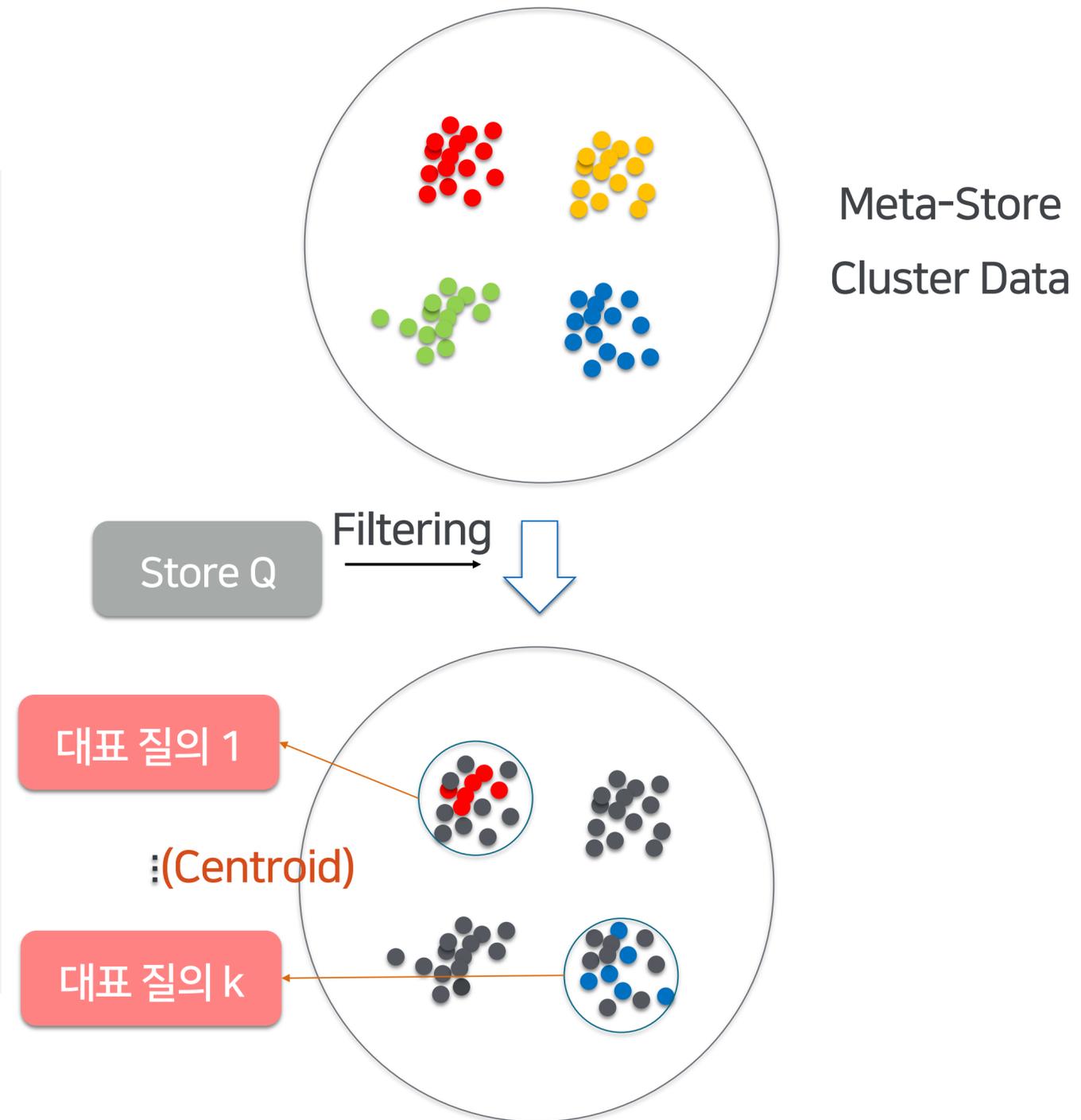
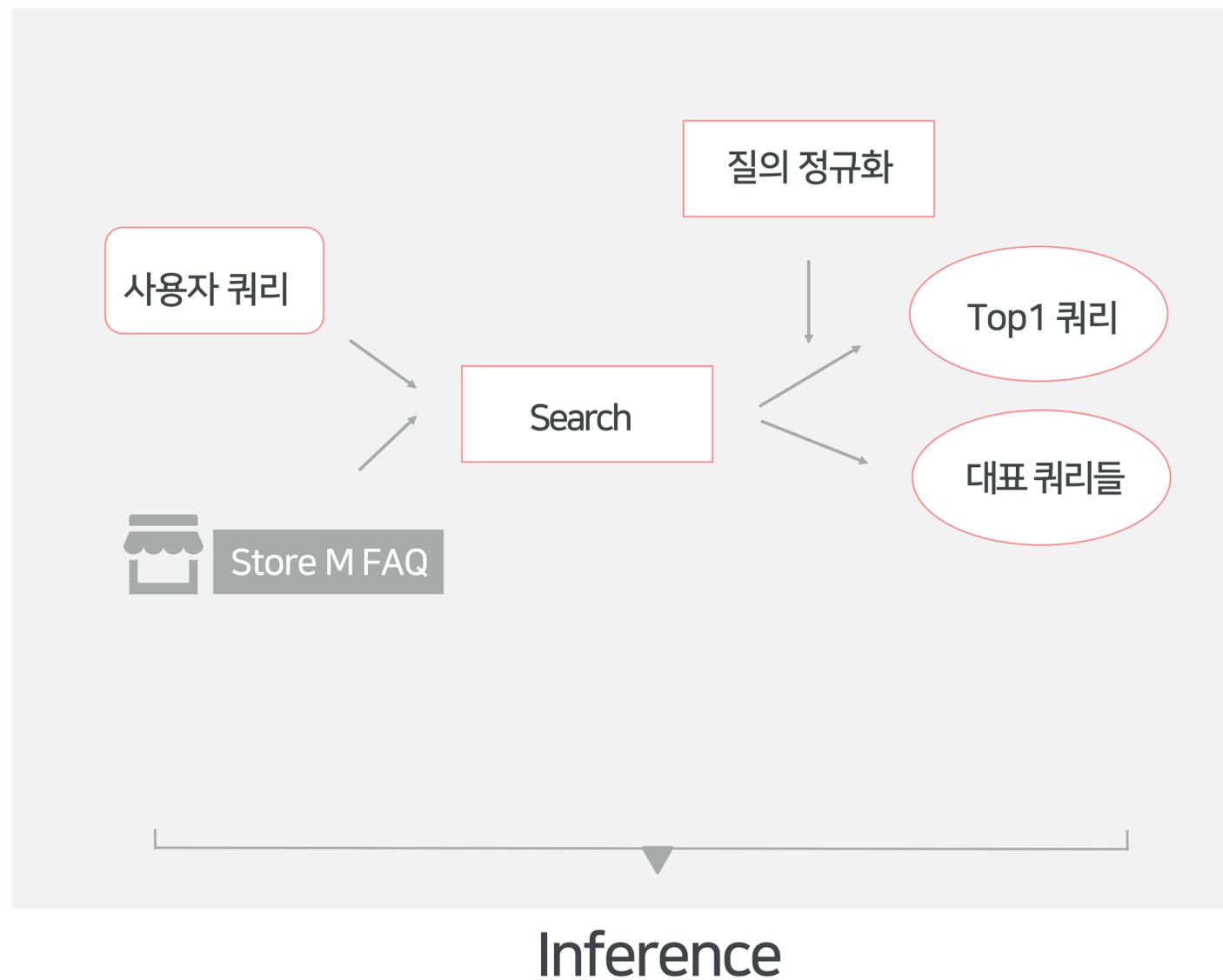
# 3.3 자동응대-FAQ: Inference

## Similarity Search

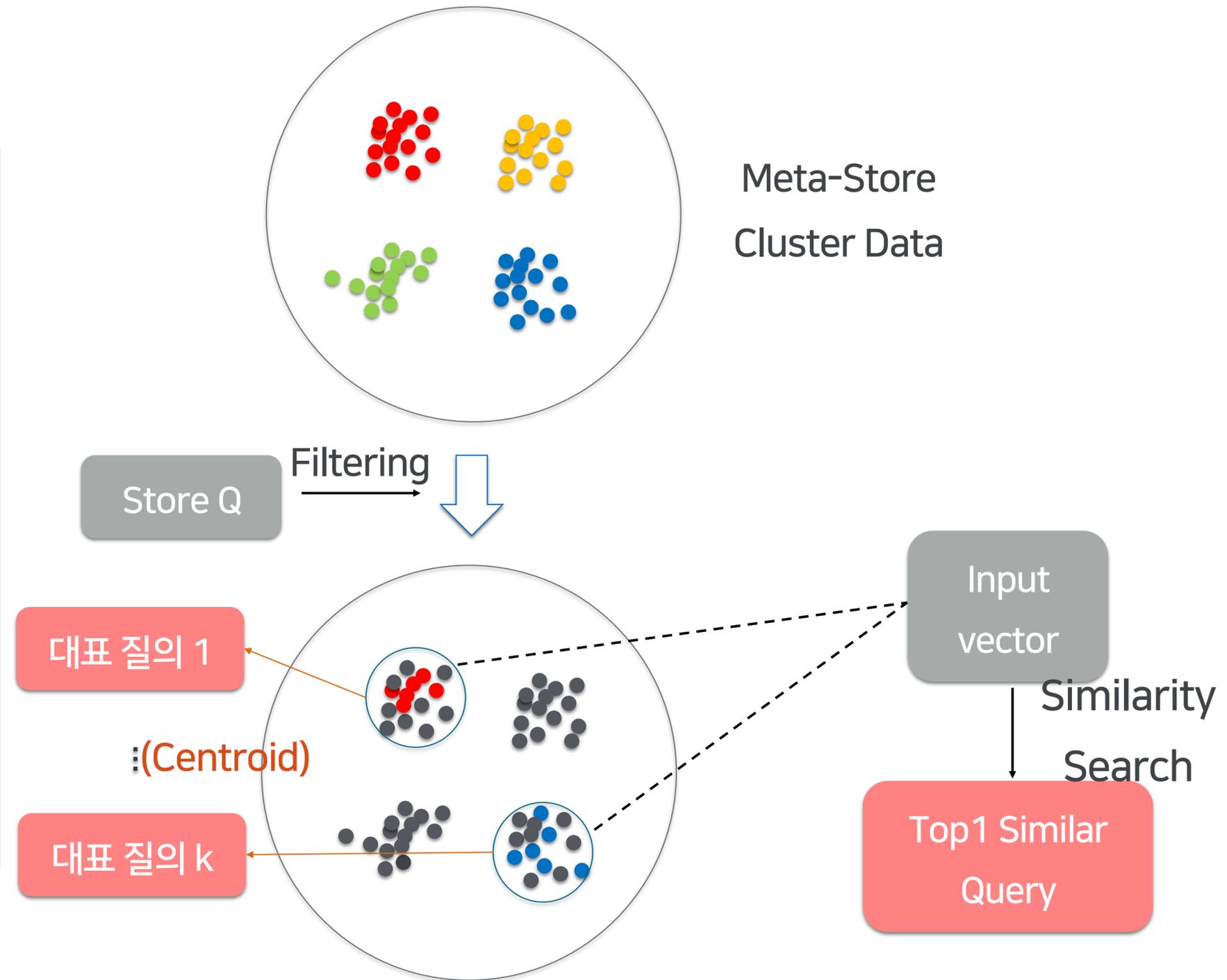
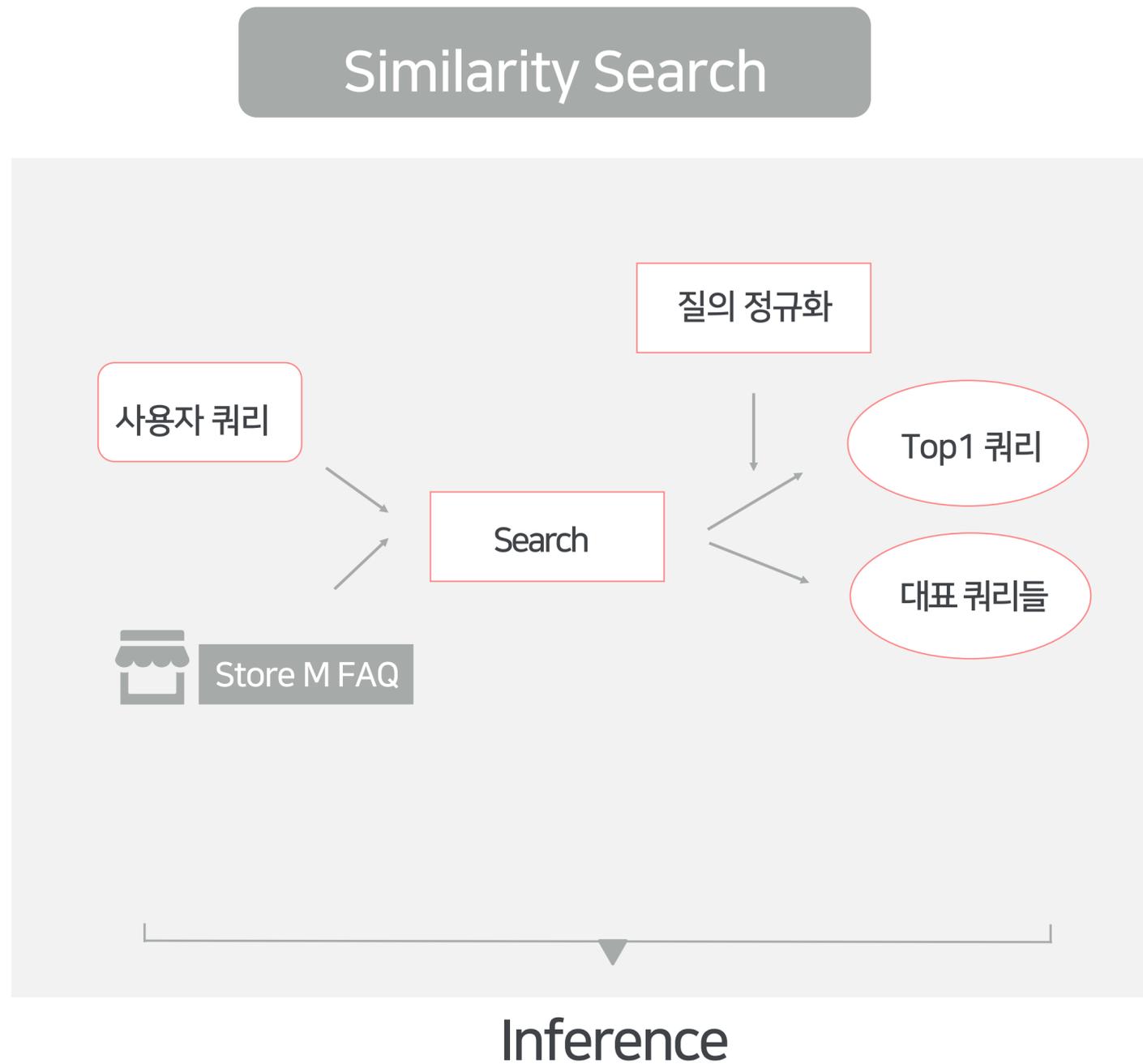


# 3.3 자동응대-FAQ: Inference

## Similarity Search



# 3.3 자동응대-FAQ: Inference



“UNICON기술 기반 FAQ서비스를 통해  
신속한 고객 맞춤형 응대 서비스 제공”

“Annotation 비용 없이 고객문의 유형을 분류하고  
최종적으로는 AI 고객센터를 만들고자 하신가요?”

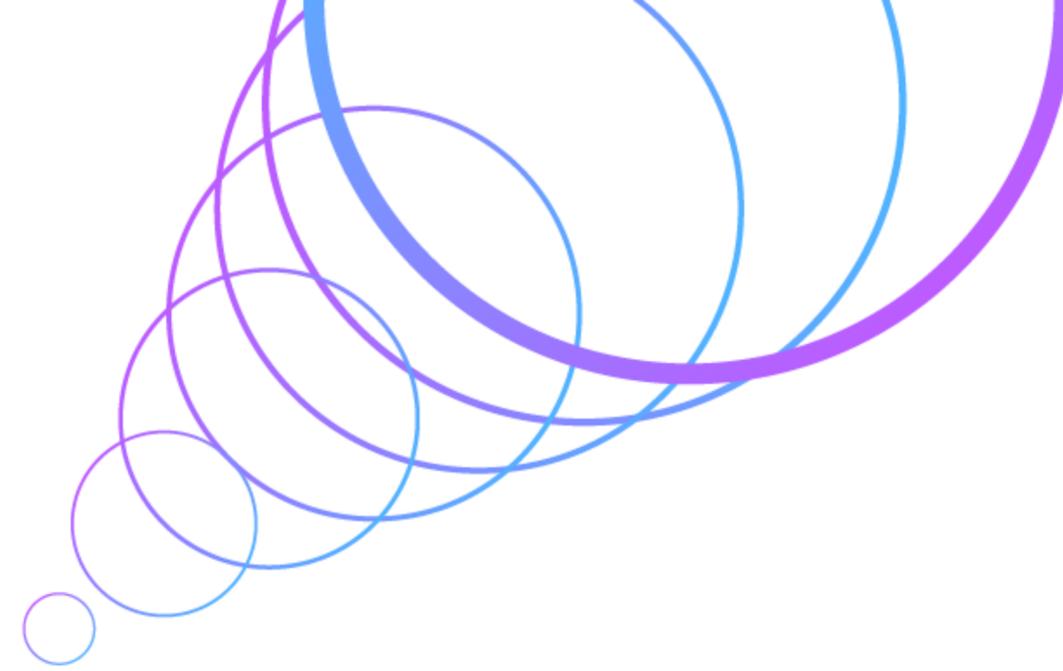
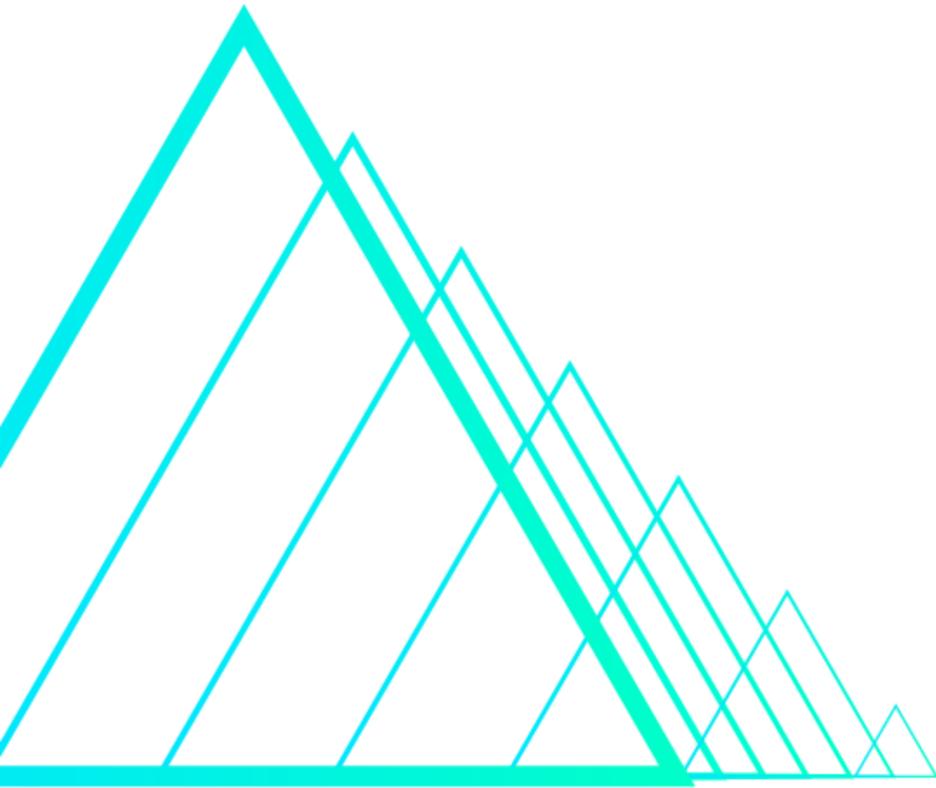
마무리

N DEVIEW  
2021

17%

“그러면 획기적으로 성능이 향상된

UNICON을 사용해 보세요!”



**Thank You**

