Multi-intent Detection

23/10/16

HYU NLP

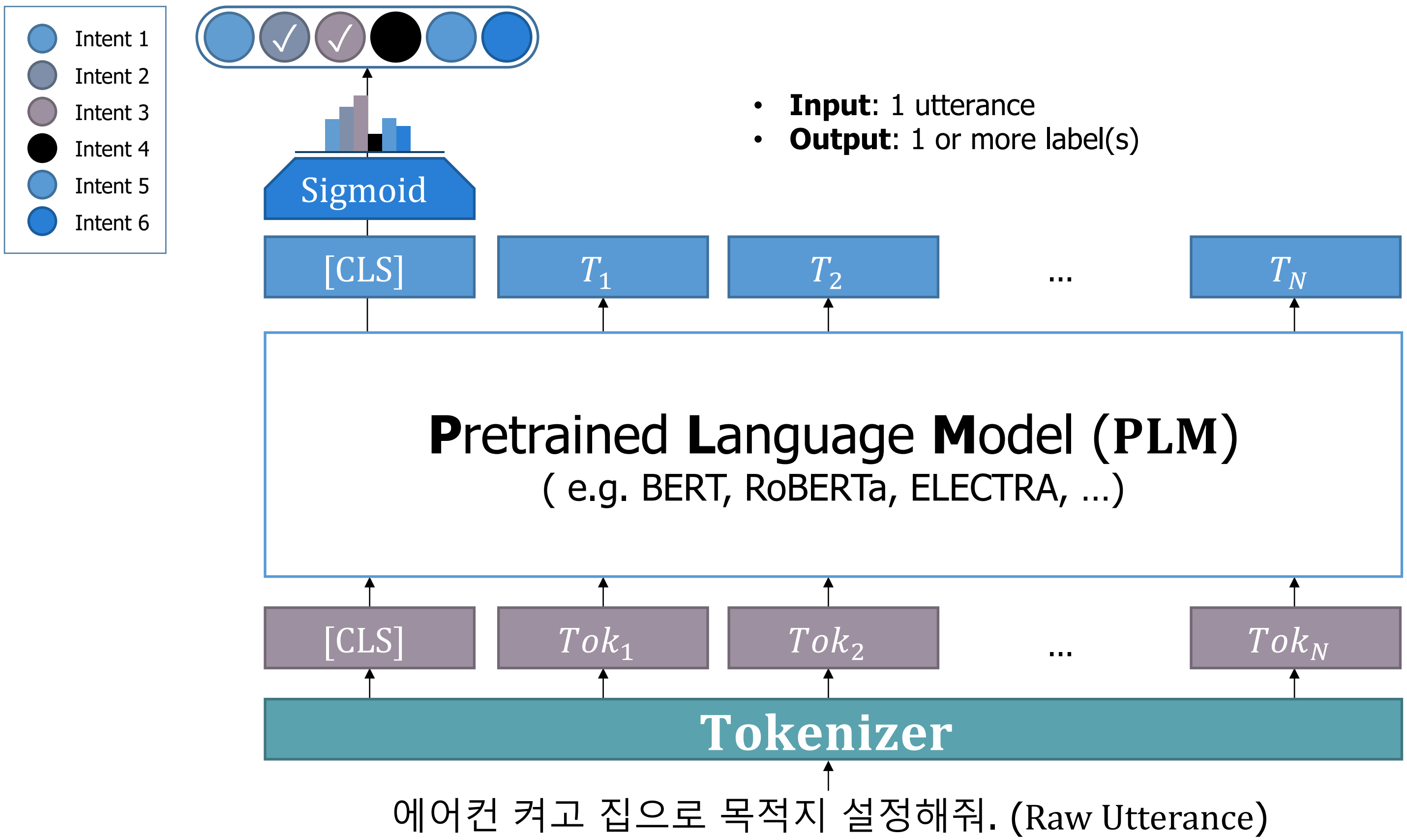
Natural Language Processing Lab.,
Hanyang University.

Contents

Task Definition

Multi-Intent Detection

- **Multi-Intent Detection**
  - 전통적인 machine learning 문제에서 정의하는 classification의 한 종류
  - 각 instance에 의도가 여러 개(multiple intent) 할당될 수 있는 classification task



- **Input**: 1 utterance
- **Output**: 1 or more label(s)

Datasets

Progress Update :: #01 Read papers on MID for survey

- **List of papers w/ categories**
  - MID, Jointly-learning (w/ Slot Filling), MLC (Multi-Label Classification), Dataset

Acronyms	Title	Authors	Released	Datasets	Categories	Date
CIBA	Multi-Point Semantic Representation for Intent Classification	Jinghan Zhang, et al.	AAAI2020	CCL,CitySrv, ECOM,TELE	MID	4/27
AGIF	AGIF: An Adaptive Graph-Interactive Framework for Joint Multiple Intent Detection and Slot Filling	Libo Qin, et al.	EMNLP2020 Findings	MixATIS, MixSNIPS	Jointly-learning	4/27
GL-GIN	GL-GIN: Fast and Accurate Non-Autoregressive Model for Joint Multiple Intent Detection and Slot Filling	Fuxuan Wei, et al.	ACL-IJCNLP 2021	MixATIS, MixSNIPS	Jointly-learning	4/27
MCT&ALR	Few-shot Learning for Multi-label Intent Detection	Yongkui Lai, et al.	AAAI2021	StandfordLU, TourSG (DSTC-4)	MID w/ dynamic threshold	4/27
SDJN	Joint Multiple Intent Detection and Slot Filling via Self-distillation	Lisong Chen, et al.	ICASSP 2022	MixATIS, MixSNIPS	Jointly-learning	4/27
ReLa-Net	Group is better than individual: Exploiting Label Topologies and Label Relations for Joint Multiple Intent Detection and Slot Filling	Bowen Xing, et al.	EMNLP2022	MixATIS, MixSNIPS	Jointly-learning	4/27
AIK	Towards Multi-label Unknown Intent Detection	Yawen Ouyang,et al.	COLING2022	MixSNIPS, MultiWOZ 2.3	MID w/ out-of-scope	4/27
HBGL	Exploiting Global and Local Hierarchies for Hierarchical Text Classification	Ting Jiang, et al.	EMNLP2022	WOS, NYT, RCV1-V2	MLC w/ label semantics	5/11
Balanced LossNLP	Balancing Methods for Multi-label Text Classification with Long-Tailed Class Distribution	Yi Huang, et al.	EMNLP2021	Reuters-21578, PubMed	MLC w/ class imbalance	5/11
MULTI-CONVFIT	Multi-Label Intent Detection via Contrastive Task Specialization of Sentence Encoders	Ivan Vulić, et al.	EMNLP2022	MixATIS, NLU++	MID w/ contrastive learning	5/11
KCOD	Watch the Neighbors: A Unified K-Nearest Neighbor Contrastive Learning Framework for OOD Intent Discovery	Yutao Mou, et al.	EMNLP2022	Banking, CLINC, HWU64	ID w/ contrastive learning	5/11
GISCo	Enhancing Joint Multiple Intent Detection and Slot Filling with Global Intent-Slot Co-occurrence	Mengxiao Song, et al.	EMNLP2022	MixATIS, MixSNIPS	Jointly-learning	5/11
DialogUSR	DialogUSR: Complex Dialogue Utterance Splitting and Reformulation for Multiple Intent Detection	Haoran Meng, et al.	EMNLP2022 Findings	*propose the datasets: DialogUSR	Dataset	5/11

Progress Update :: #04 Visualization of utterance embedding

1. Single-intent dataset의 선정

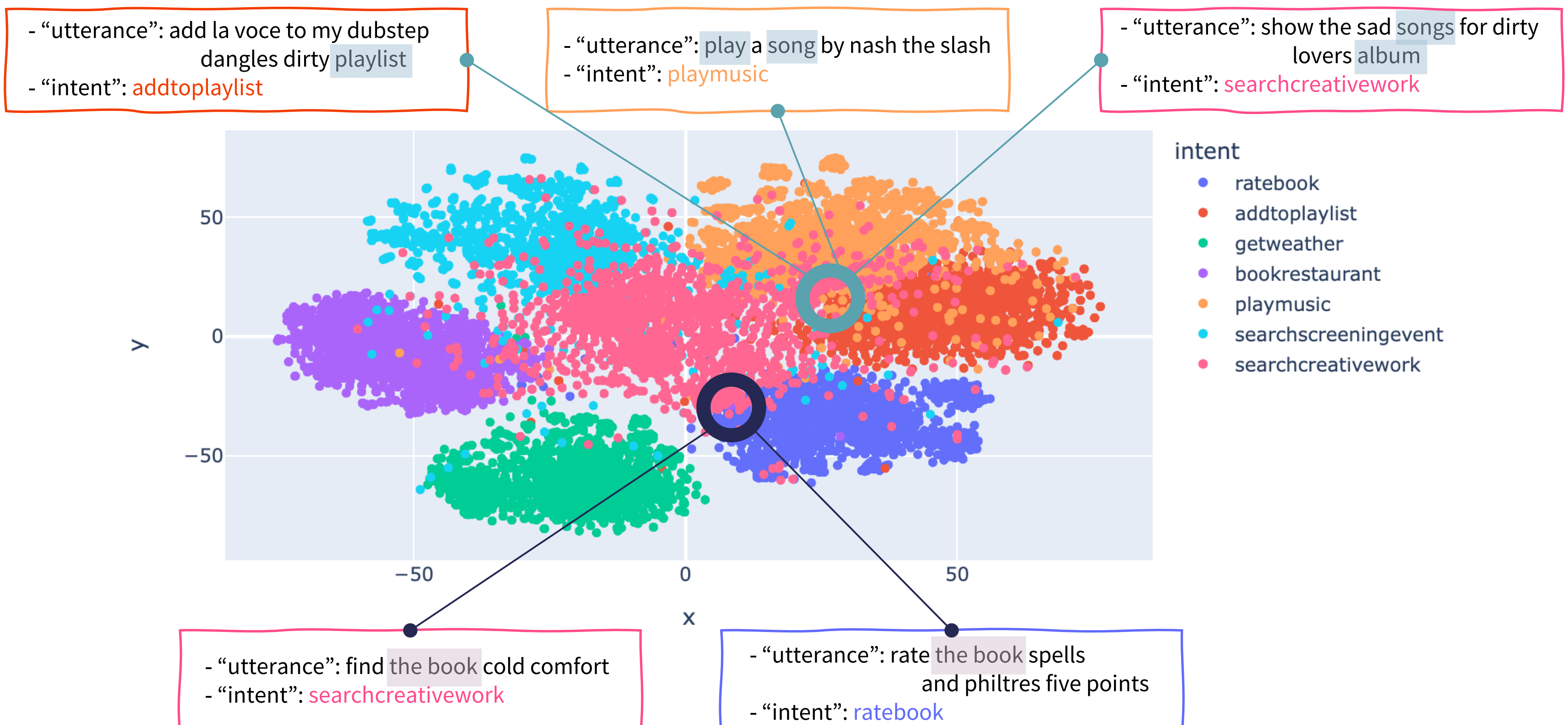
- 선정 데이터셋 : SNIPS, ATIS, Banking77, CLINC, SLURP

DATASETS	중요도	#Domain	#Intent	#Utterance	사용가능성
SNIPS	****	1	7	14,484	**High**
ATIS	****	1	22	5,871	**High**
Banking77	***	1	77	13,083	**High (Preprocessing required)**
CLINC	***	10	15×#(Domain)	23,700	**High**
HWU64	***	21	64	25,000	Mid (Ambiguous utterances)
Chatbot Corpus	**	1	2	206	Low (Too few intents)
StackExchange Corpus	**	2	12	290	Very Low (not utterance)
M2M	**	2	3	(1,500)	Low (Multi-turn)
MultiWOZ 2.2	**	8	2	(8,438)	Low (Multi-turn)
MDC	**	3	11×#(Domain)	(10,087)	Very Low (intent = Act)
SGD	**	16	86	(16,142)	Low (Multi-turn)
SLURP	*	18	77	16,521	**High**

Progress Update :: #04 Visualization of utterance embedding

2. Single-intent dataset의 시각화

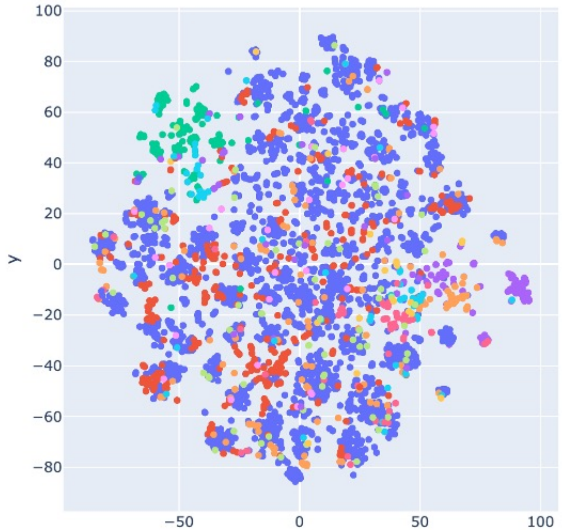
- 데이터셋 별 시각화 결과: SNIPS 예시
  - ○ 서로 다른 intent를 가지지만 embedding space 상 가까운 utterance들은 유사한 의미적, 구조적 특징을 지님



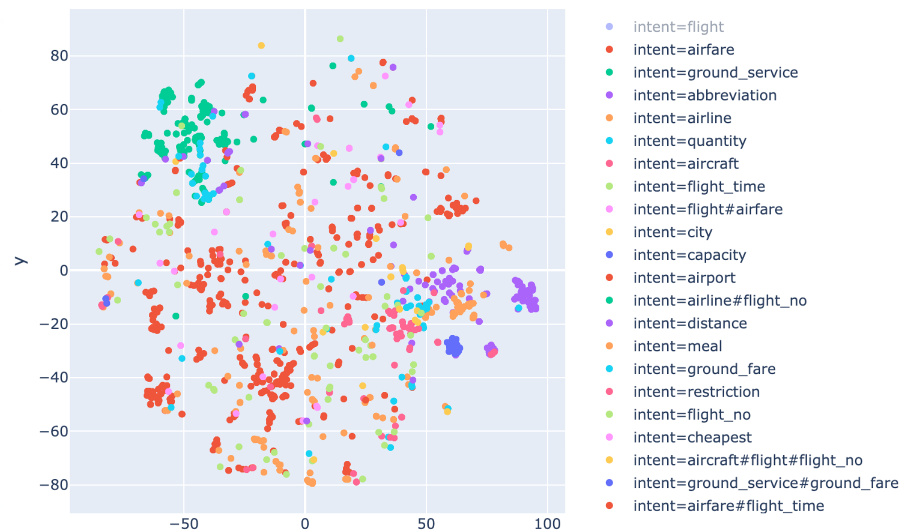
Progress Update :: #04 Visualization of utterance embedding

2. Single-intent dataset의 시각화

- 데이터셋 별 시각화 결과: ATIS 예시
  - 각 intent 별 군집 구분이 어려움
  - intent 불균형이 심한 데이터셋으로, 전체의 70%를 차지하는 intent(flight) 제외 후 상대적인 군집 확인



▲ flight 포함



▲ flight 제외

MIC #0 ~~Mix~~SNIPS

- **SNIPS: single intent classification, slot filling** (Introduced by Coucke et al. 2018)
  - One of Natural Language Understanding benchmark from Snips
  - Size: <2000 utterances over each intent
  - Label: 7 intents (from API.ai, Wit.ai, Luis.ai, Alexa and Snips)

Intent sample

- SearchCreativeWork
  Find me the I
  Robot television show
- GetWeather
  Is it windy in Boston
  MA right now?
- BookRestaurant
  I want to book a highly rated restaurant for me and my boyfriend tomorrow night
- PlayMusic
  Play the last track from Beyoncé off Spotify
- AddToPlaylist
  Add Diamonds to my roadtrip playlist
- RateBook
  Give 6 stars to Of Mice and Men
- SearchScreeningEvent
  Check the showtimes for Wonder Woman in Paris

* Repository: /sonos/nlu-benchmark/tree/master/2017-06-custom-intent-engines
* Leaderboard: paperswithcode.com/dataset/snips 🤗 snips_built_in_intents

MIC #1 MixSNIPS

- **MixSNIPS: multi-intent classification, slot filling** (Introduced by Qin et al. 2020)
  - Advanced SNIPS for multi-intent classification
  - Just concatenate sentences using "and" with different intents
    - Ratio of sentences: [1, 2, 3] intents [0.3, 0.5, 0.2]
  - Size: 50,000 utterances
  - Label: 7 intents (up to 3-label)

* Repository: [/LooperXX/AGIF]()

* ~~Leaderboard~~

- Intent Detection: 97.7 (accuracy)
- Slot Filling: 96.4 (F1)

sample ▼

BookRestaurant

```
book O
a O
reservation O
for O
my B-party_size_description
mommy I-party_size_description
and I-party_size_description
i I-party_size_description
at O
a O
restaurant B-restaurant_type
in O
central B-country
african I-country
republic I-country
and O
then O
play O
the O
newest B-sort
melody B-music_item
on O
last B-service
fm I-service
by O
eddie B-artist
vinson I-artist
BookRestaurant#PlayMusic
```

PlayMusic

* intents (indicator: '#')

MIC #2 MixATIS

- **MixATIS: multi-intent classification, slot filling** (Introduced by Qin et al. 2020)
  - Advanced ATIS for multi-intent classification
  - Just concatenate sentences using "and" with different intents
    - Ratio of sentences: [1, 2, 3] intents [0.3, 0.5, 0.2]
  - Size: 20,000 utterances
  - Label: 22 intents (up to 3-label)

* Repository: [/LooperXX/AGIF]()

* Leaderboard: [paperswithcode.com/dataset/mixatis]()

  - Intent Detection: 76.3 (accuracy)
  - Slot Filling: 88.3 (F1)
  - Semantic Frame Parsing: 43.5 (accuracy)

sample

atis_capacity

```
how O
many O
passengers O
can O
a O
boeing O
737 B-aircraft_code
hold O
and O
```

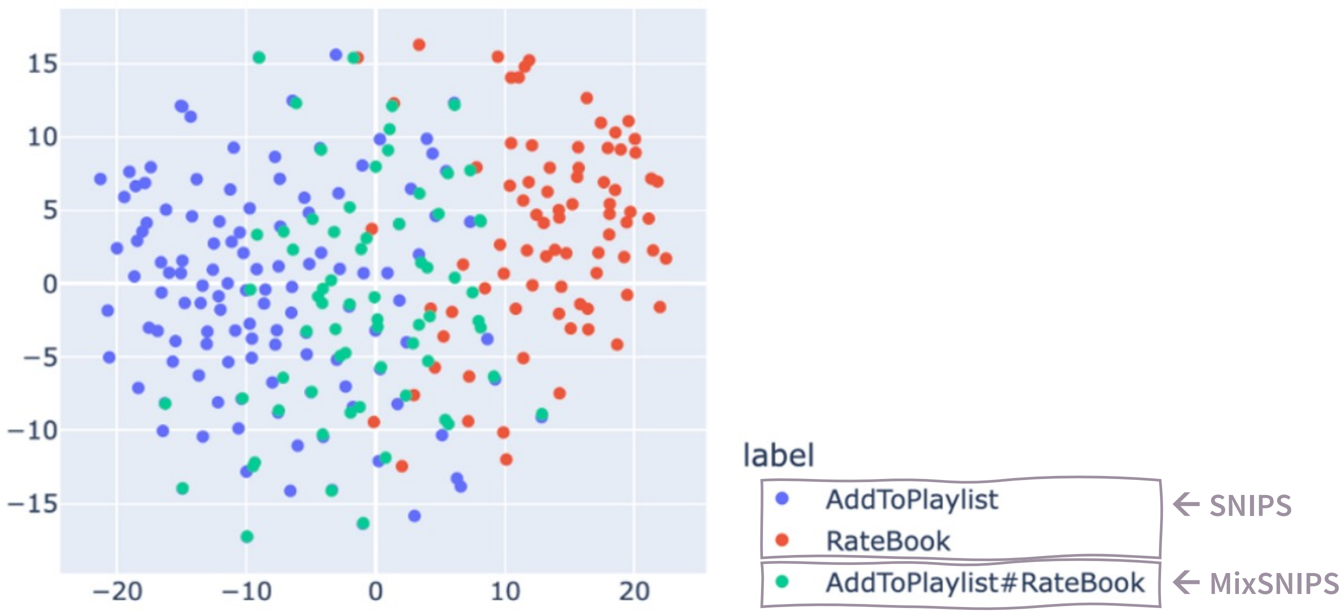
atis_flight

```
also O
flights O
from O
pittsburgh B-fromloc.city_name
to O
baltimore B-toloc.city_name
between O
10 B-depart_time.start_time
am I-depart_time.start_time
and O
2 B-depart_time.end_time
pm I-depart_time.end_time
atis_capacity#atis_flight
```

* intents (indicator: '#')

2. Single-intent dataset의 시각화

- multi-intent utterance와 single-intent utterance 간 embedding 비교
  MixSNIPS SNIPS

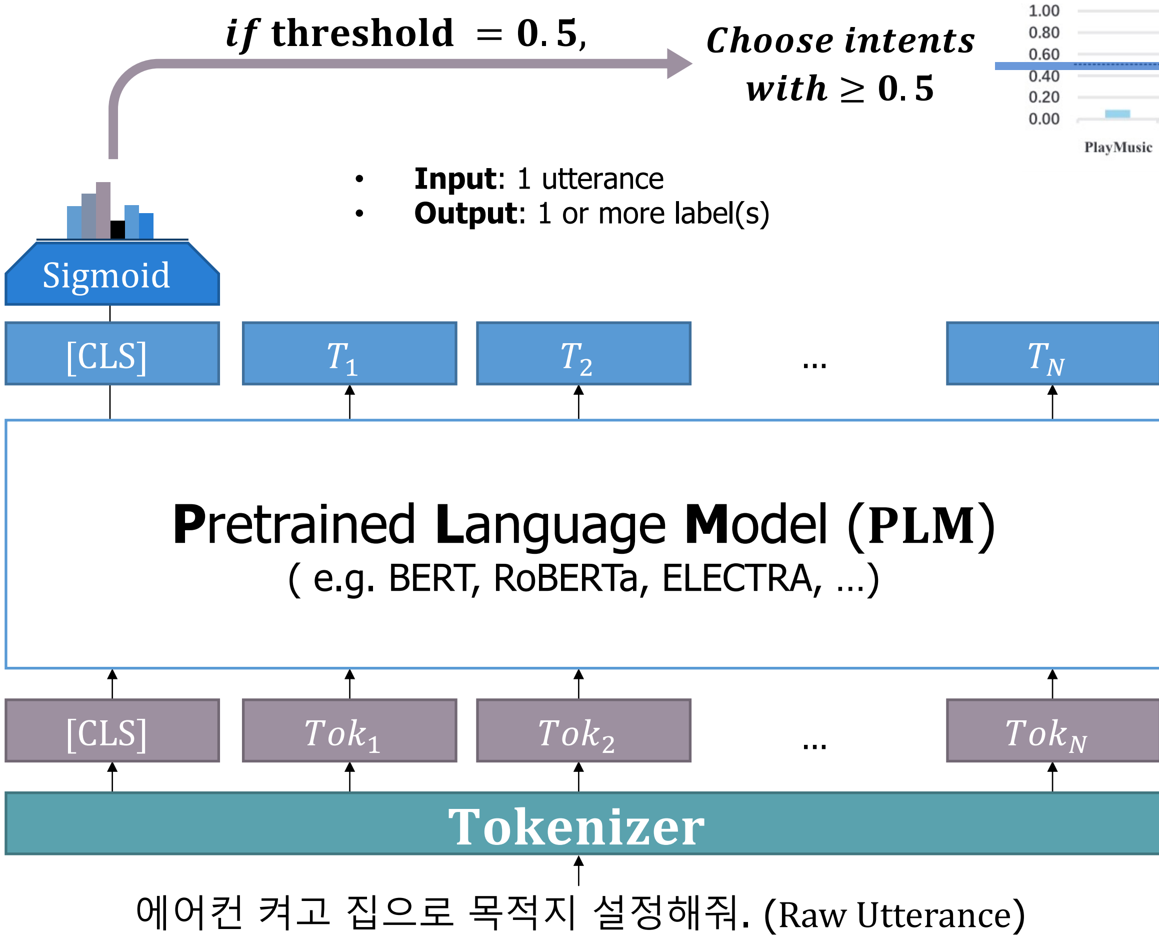


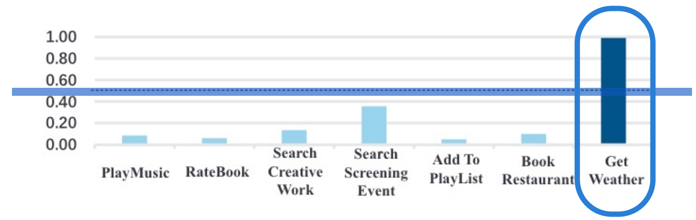
→ multi-intent인 AddToPlaylist#RateBook 군집이
single-intent인 AddToPlaylist 군집과 RateBook 군집 사이에 위치함

Approach

MLC with Threshold

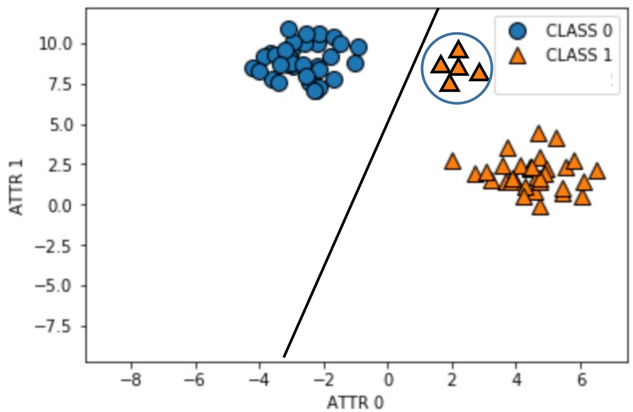
- **Apply Sigmoid for multi-label classification problem**
  - classification threshold은 default로 0.5로 선정

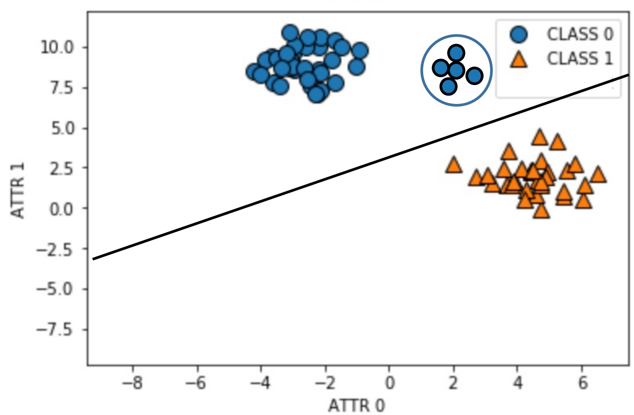


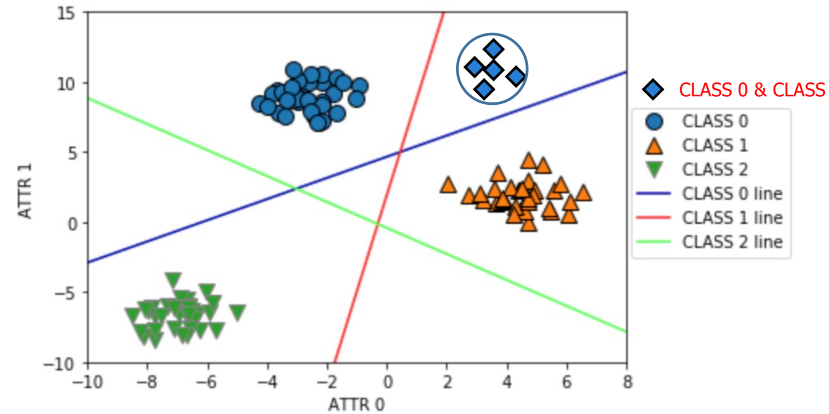


MLC without Threshold – OvR

- **Use One-vs-Rest Method**
  - A heuristic method for using <u>binary classification</u> for multi-class classification
  - Label 수 만큼의 이진 분류 모델을 학습, 1회의 추론에서 모든 이진 분류 모델을 사용



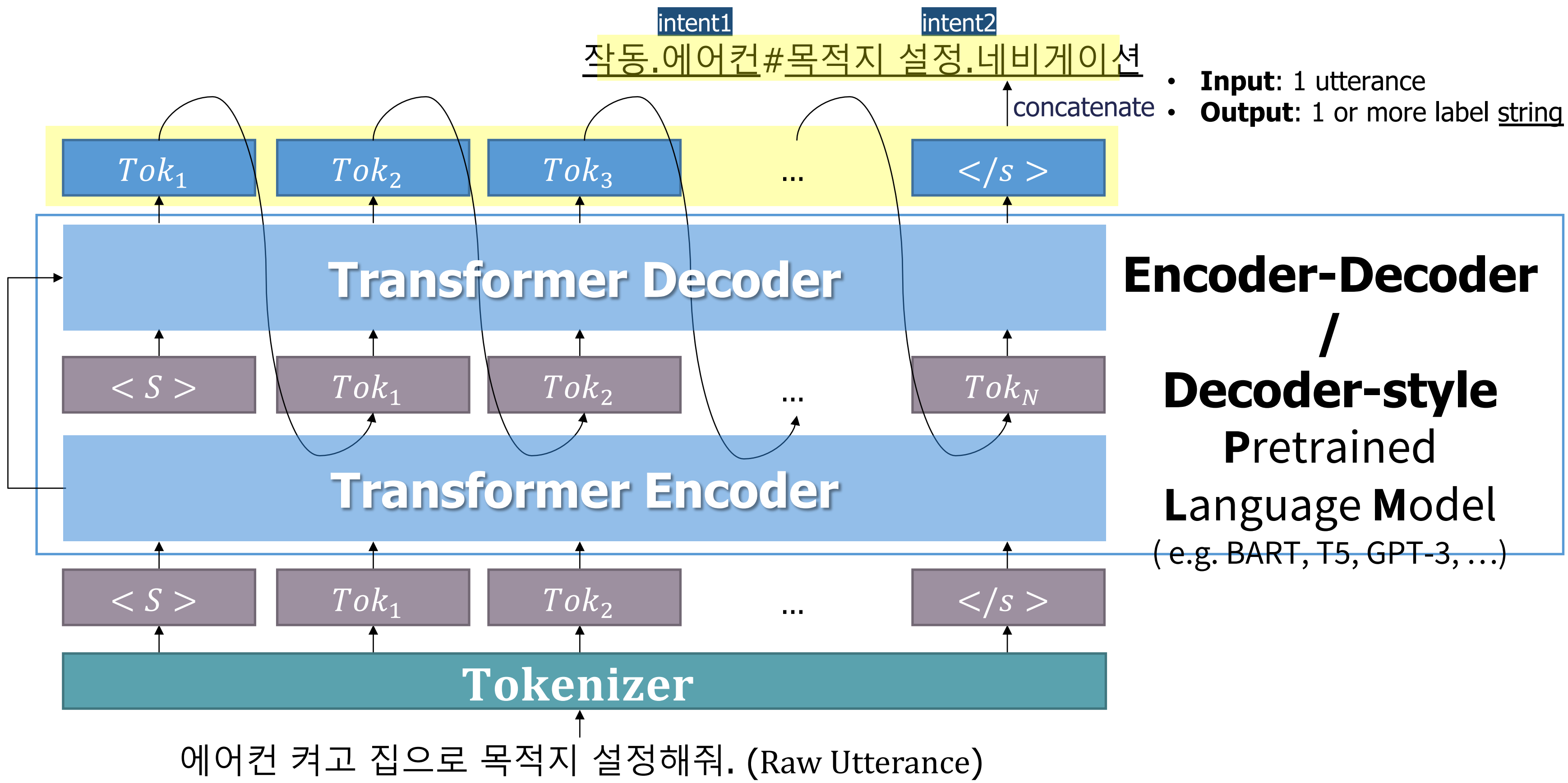




Multi-label classification **using OvR**

Generative Classification

- **Apply *text-to-text* framework**
  - Decoder-style 모델이 input으로 받은 utterance에 대해 output으로 분류할 label 조합을 출력하도록 transfer learning



Baseline

MLC without Threshold – jointly-learning

TFMN) Chen et al. A Transformer-based Threshold-Free Framework for Multi-Intent NLU, COLING, 2022

* Multi-task learning: #01 + #02 + #03

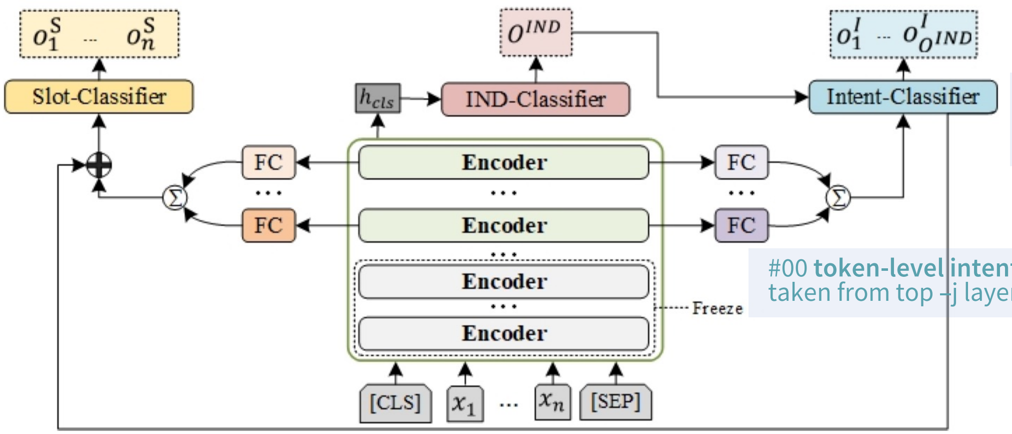
- **Suggestion**

#01 **IND cls**: classify # of intent

#03 **Slot cls:** seq labeling

#02 **Intent** MLC: classify top-k multi-intent based on prediction of **IND**

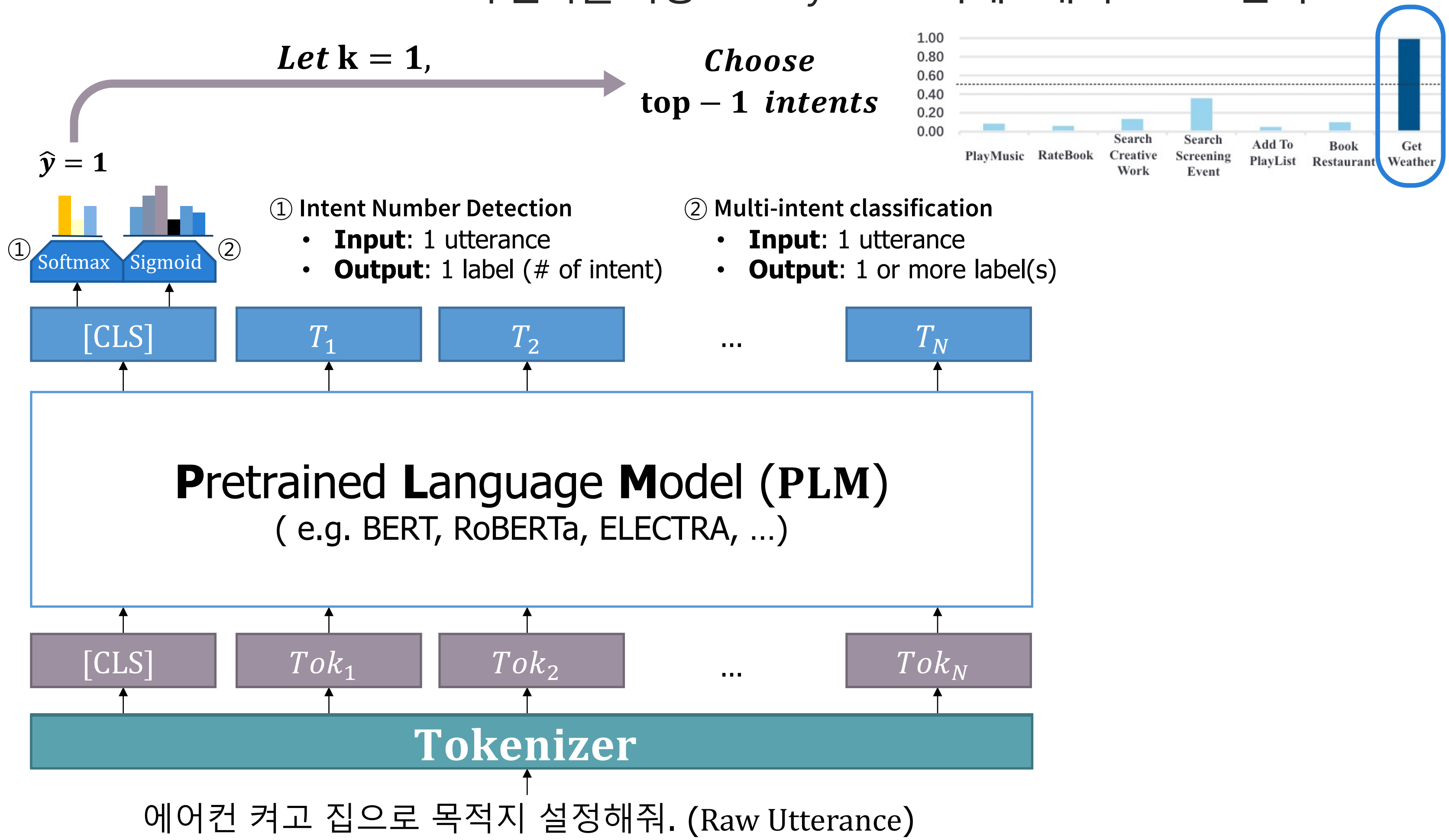
#00 **token-level intent feature** taken from top -j layer of encoder



- Transformer-based thresholdless multi-intent NLU framework w/ 3 multi-task learning – Intent cls, IND cls*, Sot cls
  * IND task: an auxiliary task, model detects the number of intents in each utterance
- The output of each upper j-layer in the encoder is used to <u>generate multi-grain representations</u> at different levels of granularity (passed through FC and just sum them up)

MLC without Threshold – jointly-learning

- **Classify top-k multi-intent based on prediction of IND**
  - Intent Number Detector의 결과를 바탕으로 Dynamic하게 K개의 intent 선택



Progress Update :: #03 Test MID w/OpenAI API

Action Item #03 Test MID w/ChatGPT

- **Result Table**

	model	dataset	metric	score	set.	note
1	TFMN	MixSNIPS	Accuracy	**0.977**		예측된 multi-intent 순서 고려
2	text-davinci-003	MixSNIPS	Accuracy	0.7312	zero	예측된 multi-intent 순서 무시
3	text-davinci-003	MixSNIPS	Accuracy	0.546	few	예측된 multi-intent 순서 무시
4	gpt-3.5-turbo	MixSNIPS	Accuracy	**0.8208**	zero	예측된 multi-intent 순서 무시
5	gpt-3.5-turbo	MixSNIPS	Accuracy	0.8508	zero	예측된 multi-intent 순서 무시 * accuracy 기준 완화 (포함만 되더라도)
6	gpt-3.5-turbo	MixSNIPS	Accuracy	TBD	few	예측된 multi-intent 순서 무시

	model	dataset	metric	score	set.	note
7	TFMN	MixATIS	Accuracy	**0.798**		예측된 multi-intent 순서 고려
8	text-davinci-003	MixATIS	Accuracy	0.327	zero	예측된 multi-intent 순서 무시
9	text-davinci-003	MixATIS	Accuracy	0.268	few	예측된 multi-intent 순서 무시

- Metric: overall accuracy
  * accuracy 기준 완화: 정답 intent + 그 외 intent를 추가로 detect한 경우 정답으로 처리
- Model: text-davinci-003, gpt-3.5-turbo

Progress Update :: #03 Test MID w/OpenAI API

☑ Action Item #03 Test MID w/ChatGPT

- **Analysis of Results**

	model	dataset	metric	score	set.
1	TFMN	MixSNIPS	Accuracy	**0.977**	
2	text-davinci-003	MixSNIPS	Accuracy	0.7312	zero
3	text-davinci-003	MixSNIPS	Accuracy	0.546	few
4	gpt-3.5-turbo	MixSNIPS	Accuracy	**0.8208**	zero
5	gpt-3.5-turbo	MixSNIPS	Accuracy	0.8508	zero
6	gpt-3.5-turbo	MixSNIPS	Accuracy	TBD	few

	model	dataset	metric	score	set.
7	TFMN	MixATIS	Accuracy	**0.798**	
8	text-davinci-003	MixATIS	Accuracy	0.327	zero
9	text-davinci-003	MixATIS	Accuracy	0.268	few

note

#(intent)	MixSNIPS #(utterance)		
	answer	few-shot	zero-shot
1	750	281	727
2	1,250	1,794	1,483
3	500	365	288
4		34	2
5		22	
6		2	
7		2	
total	2,500	2,500	2,500

- text-davinci-003 기준, 모든 데이터셋에서 Zero-shot 보다 Few-shot 성능이 낮음
  - Zero-shot 대비 Few-shot에서 SID(Single-intent Detection)에 대해 큰 폭의 성능 저하
    - single-intent utterance 750개 중, 60% 이상을 multi-intent로 detect함 (MixSNIPS)

Contents

- **Single-intent to Multi-intent**
  - Single-intent Datasets
  - Multi-intent Datasets
- **Multi-intent Detection SOTA baselines**
  - TFMN
  - SLIM
  - ChatGPT
- **Concatenation –side**
- **Selection -side**

Our Approach

Abstract

- **데이터 구축 관점에서의 Multi-intent setting**
  - 기존의 naïve 하게 병합된 multi-intent utterance setting을 연구적으로 고도화하고 현실적인 문제로 확장

  → 모델이 (혹은 인간이) 판단하기에 충분히 어렵고 미묘하게 병합된 multi-intent utterance 데이터셋 공개

MixSNIPS MixATIS

Multi-intent Datasets

- play an album by roni duani and also can i get the game list of mew singles
- how much is limousine service in los angeles and also is there ground transportation available at the denver airport

ATIS Banking77 SNIPS CLNC150 SLURP HWU64

Single-intent Datasets

Naïve Approach

- 상관성을 고려하지 않은 **대등관계** 문장
- **Template-based** Concatenation
  - **접속사** 활용: and then, and, also, next, for, yet, however, nevertheless, after, because, than, where, while, if, unless, first of all, finally, I also would like to know, additionally, …
- Sentence Selection
  : Simple Random Sampling
- Intent 조합
  : 의미적 연관성 낮음
  : 동시 발생 가능성도 낮음

* MixSNIPS, MixATIS setting

Soft

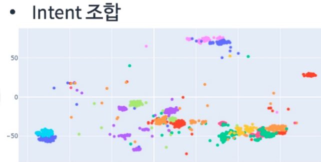
Complex Approach

- 보다 상관성 있는 **대등관계** 문장
- **Template-based** Concatenation
  - **접속사** 활용: (좌동)
- Sentence Selection
  : **Utterance Representation** 활용
  1) Encode Utterance: Vector space에 intent별로 clustering되도록 utterance embedding 학습
  2) 서로 다른 intent이지만 utterance간 거리가 가까운 2개 utterance 선택
  3) Template-based merge → multi-intent utterance 생성
- Intent 조합
  : 의미적 연관성 높음
  : 동시 발생 가능성 모호

Hard

Complex Approach

- 상관성을 고려한 **종속관계** 문장
- **생략, 치환, 상호참조, 순서** 등 활용
  - 동일한 주어/동사 생략
  - 대명사, 지시어 등 사용
  - 순서가 있는 intent 조합
  - …
- Sentence Selection
  : 연구적 고민 필요
- Intent 조합



Discussion

• 데이터 구축 관점에서의 Multi-intent setting

Soft

Hard

Naïve Approach

- 상관성을 고려하지 않은 **대등관계** 문장
- **Template-based** Concatenation
  - **접속사** 활용: and then, and, also, next, for, yet, however, nevertheless, after, because, than, where, while, if, unless, first of all, finally, I also would like to know, additionally, …
- Sentence Selection
  : Simple Random Sampling
- Intent 조합
  : 의미적 연관성 낮음
  : 동시 발생 가능성도 낮음

* MixSNIPS, MixATIS setting

e.g. A음악 켜고 B영화 켜줘
1) A음악 켜줘 (#music.play:A음악)
2) B영화 켜줘 (#movie.play:B영화)

Complex Approach (Soft)

- 보다 상관성 있는 **대등관계** 문장
- **Template-based** Concatenation
  - **접속사** 활용: (좌동)
- Sentence Selection
  : **Utterance Representation** 활용
  1) Encode Utterance: Vector space에 intent별로 clustering되도록 utterance embedding 학습
  2) 서로 다른 intent이지만 utterance간 거리가 가까운 2개 utterance 선택
  3) Template-based merge → multi-intent utterance 생성
- Intent 조합
  : 의미적 연관성 높음
  : 동시 발생 가능성 모호

e.g. A음악 켜고
A음악 플레이리스트에 넣어줘
1) A음악 켜줘 (#music.play:A음악)
2) A음악 플레이리스트에 넣어줘
(#music.add_playlist:A음악)

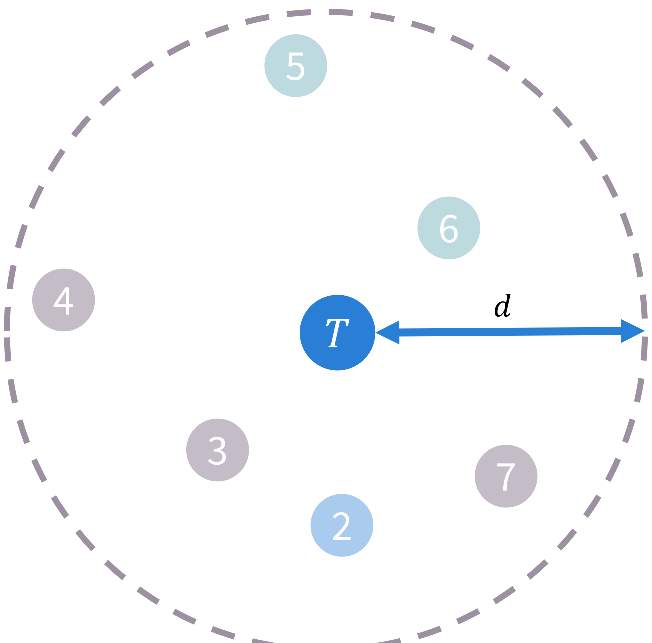
Complex Approach (Hard)

- 상관성을 고려한 **종속관계** 문장
- **생략, 치환, 상호참조, 순서** 등 활용
  - 동일한 주어/동사 생략
  - 대명사, 지시어 등 사용
  - 순서가 있는 intent 조합
  - …
- Sentence Selection
  : 연구적 고민 필요
- Intent 조합
  : 의미적 연관성(종속성) 높음
  : 동시 발생 가능성도 높음

e.g. A음악 켜고
플레이리스트에 넣어줘
1) A음악 켜줘 (#music.play:A음악)
2) ~~A음악~~ 플레이리스트에 넣어줘
(#music.add_playlist:A음악)

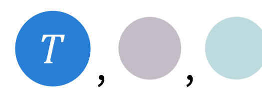
1-1. Embedding-based utterance selection to merge

- 절차
  6. 4번의 결과로 나온 target utterance와 candidate로 만들 수 있는 조합들 중,
     무작위 샘플링 하여 utterance들을 template 등을 활용하여 concatenation할 예정



원 내부 숫자
: 각 embedding의 index

e.g.) intent가 3개인 multi-intent utterance를
생성하는 가능한 조합



: $(T, 3, 5)$, $(T, 3, 6)$,
$(T, 4, 5)$, $(T, 4, 6)$,
$(T, 7, 5)$, $(T, 7, 6)$

Random sample

$(T, 4, 6)$

최종 multi-intent utterance:
***T* and 4 and then 6**

Progress Update

Action Item #09 Combine single-intent utterances to a multi-intent utterance

- **Concatenation**

Options		Description	Example
Template-based	AGIF setting	MixSNIPS, MixATIS 사용한 conjunctions	‘,’, and, and then, and also
	various conj. pool	성능 확인 목적 실험에는 당초 구성 예정의 conjunction 중 일부만 활용	or, before, after, additionally, finally
Template-free	removal	연결어가 아예 없이 두 발화가 연속해서 들어온 경우 가정	
	removal + gerunds	연결어구 없이 발화를 잇는 자연스럽고 발생가능한 상황 가정	
ChatGPT-based		Cosine sim. 0.7 이상인 2개 문장을 gpt-3.5-turbo로 연결	* 다음 슬라이드 예시

Progress Update

Action Item #09 Combine single-intent utterances to a multi-intent utterance

- **Concatenation**
  - Cherry🍒

what are the flights between washington dc and columbus ohio	**atis_flight**	0.9227	What flights are available between Washington DC and Columbus Ohio and which airlines operate them?
what airlines fly between washington dc and columbus ohio	**atis_airline**		

what are the flights from cleveland to dallas	**atis_flight**	0.8041	What are the flights from Cleveland to Dallas, and what type of aircraft are flying before noon?
what type of aircraft are flying from cleveland to dallas before noon	**atis_aircraft**		

  - Lemon🍋

What are the fees for top-ups?	**top_up_by_bank_transfer_charge**	0.9813	What are the fees for top-ups, whether by bank transfer or by card?
What are the fees for top ups?	**top_up_by_card_charge**		

what are the flights from milwaukee to seattle	**atis_flight**	0.7230	What are the flight numbers from Milwaukee to Seattle and Chicago to Seattle?
flight numbers from chicago to seattle	**atis_flight_no**		

Experiment

- **Ours vs mix- setting**

2nd)	23.10.16						
Model		Option		Dataset (accuracy)			
		training	test	SNIPS	ATIS	Banking77	CLINC150
supervised	tfmn-sf	mix-setting	mix-setting	95.96	76.80	76.11	85.60
		mix-setting	our setting	51.01	50.40	36.96	46.15
		our setting	our setting	92.96	76.00	62.69	78.06
	slim-sf	mix-setting	mix-setting	95.88	91.48	0.06	85.85
		mix-setting	our setting	92.96	64.09	0.06	74.47
		our setting	our setting	95.72	77.33	0.10	84.44
unsupervised	gpt-3.5-turbo	-	mix-setting	77.56	33.60		
		-	our setting	73.23	29.96		40.98

Discussion

- **데이터 구축 관점에서의 Multi-intent setting**

	fail			~~21%~~	16%		~~7%~~	14%
		SNIPS n=91						
		A	B	C	C'	D	E	E'
Length 유의점: ≤0	value	LENGTH(0 이하가 유의하게 증가하는 경향(banking77 제외), 4 이하 평균 역시 감소)						
	-11							
	-10							
	-9							
	-8							
	-7							
	-6						1	1
	-5						7	7
	-4							
	-3						4	4
	-2						8	7
	-1		1	10	7		6	6
	0		36	19	16	31	12	10
	1	29	45	31	30	47	29	28
	2	71	18	16	14	13	9	7
	3			8	7		5	3
	4			6	6		3	1
	5			4	2		2	1
	6			2	1		1	1
	7						1	1
	8			2	1		1	1
	9							
	10						1	
	11			1				
	19			1				
	25						1	
	0이하	0	37	29	23	31	38	35
	1~4	100	63	61	57	60	46	39
	5이상	0	0	10	4	0	7	4
	0이하/4이하	0%	59%	48%	40%	52%	83%	90%
	4이하mean	1.7100	0.8000	1.1222	1.2000	0.8022	-0.0119	-0.2432
	SUM	100	100	100	84	91	91	78
	mean	1.71	0.818182	1.79	1.43	0.8022	0.7143	0.1026
Conj. 유의점: ≤0		AND + Conjunction						
	-1							
	0		56	10	7	45	3	1
	1	100	36	58	49	37	72	66
	2		8	24	22	9	13	10
	3			8	6		3	1
	4							
	0미만	0	0	0	0	0	0.000	0.000
	0이하	0	56	10	7	45	3	1
	SUM	100	100	100	84	91	91.000	78.000
	mean	1	0.52	1.3	1.321429	0.6044	1.176	1.141
		AND (ChatGPT는 AGIF의 naive concatenation_and 와 다른 방식을 사용했을 것으로 기대) > mean 경향상 ChatGPT는 concat에 and를 AGIF보다 더! 사용함						
	-1							
	0		88	38	25	78	12	7
	1	100	12	62	59	13	78	71
	2						1	
	0이하	0	88	38	25	78	12	7
	SUM	100	100	100	84	91	91	78
	mean	1.00	0.12	0.62	0.70	0.1429	0.8791	0.9103
		Conjunction(ChatGPT는 and 이외의 연결사를 더 많이 사용하거나 연결사를 아예 사용하지 않을 것으로 기대)						
	-1							
	0	100	68	45	41	58	70	63
	1		24	44	35	24	15	12
	2		8	9	7	9	6	3
	3			2	1			
	0이하	100	68	45	41	58	70	63
	SUM	100	100	100	84	91	91	78
	mean	0.00	0.4	0.68	0.62	0.4615	0.2967	0.2308
Pronoun 유의점: ≥1		Pronoun (ChatGPT가 생략을 잘 했다면 대명사를 사용했을 것이다) > mean 경향 부합						
	-3							
	-2			1			1	1
	-1			2	2		2	1
	0	100	100	93	81	91	66	70
	1			4	1		22	6
	2							
	0미만(생략)	0	0	3	2	0	3	2
	1이상	0	0	4	1	0	22	6

	fail			~~22%~~	41%		~~10%~~	10%
		ATIS n=100						
		A	B	C	C'	D	E	E'
Length	-11						1	
	-10						1	1
	-9						2	
	-8						1	
	-7						1	
	-6						3	1
	-5						2	1
	-4						4	3
	-3			1				
	-2						3	2
	-1			4			6	6
	0		36	13	6	33	11	11
	1	32	43	39	34	52	32	32
	2	68	21	13	10	15	7	7
	3			13	5		14	14
	4			4	2		5	5
	5			5	2		4	4
	6			5			2	2
	7							
	8			1			1	1
	9			1				
	10			1				
	11							
	19							
	25							
	0이하	0	36	18	6	33	35	25
	1~4	100	64	69	51	67	58	58
	5이상	0	0	13	2	0	7	7
	0이하/4이하	0%	56%	26%	12%	49%	60%	43%
	4이하mean	1.6800	0.8500	1.2989	1.3509	0.8200	-0.0215	0.7831
	SUM	100	100	100	59	100	100	90
	mean	1.68	0.85	1.95	1.4746	0.82	0.38	1.1667
Conj. (AND + Conjunction)	-1						3	1
	0		52	15	2	50	9	3
	1	100	40	35	21	40	49	47
	2		8	42	32	10	28	28
	3			7	3		5	5
	4			1	1		6	6
	0미만	0	0	0	0	0	3	1
	0이하	0	52	15	2	50	12	4
	SUM	100	100	100	59	100	100	90
	mean	1	0.56	1.44	1.6610	0.6	1.41	1.5667
(AND)	-1						2	1
	0		85	29	7	86	21	12
	1	100	15	70	51	14	77	77
	2			1	1			
	0이하	0	85	29	7	86	23	13
	SUM	100	100	100	59	100	100	90
	mean	1.00	0.15	0.72	0.90	0.14	0.75	0.84
(Conjunction)	-1						2	1
	0	100	67	38	19	64	49	42
	1		25	54	36	26	37	35
	2		8	6	3	10	5	5
	3			2	1		7	7
	0이하	100	67	38	19	64	51	43
	SUM	100	100	100	59	100	100	90
	mean	0.00	0.41	0.72	0.76	0.46	0.66	0.72
Pronoun	-3							
	-2							
	-1						3	1
	0	100	100	92	58	100	86	79
	1			8	1		10	9
	2						1	1
	0미만(생략)	0	0	0	0	0	3	1
	1이상	0	0	8	1	0	11	10

	fail			~~15%~~	22%		~~6%~~	9%
		Banking77 n=100						
		A	B	C	C'	D	E	E'
Length	-11							
	-10			1				
	-9							
	-8						1	1
	-7							
	-6			1	1			
	-5			1				
	-4			1			2	2
	-3			2	2		7	7
	-2		2	3	2	4	4	4
	-1		10	8	5	2	6	6
	0		34	20	17	33	9	8
	1	35	44	45	36	52	46	43
	2	64	9	11	9	9	9	7
	3	1		5	5		5	5
	4			2	1		3	3
	5						2	1
	6		1				3	3
	7						2	
	8							
	9							
	10							
	11						1	1
	19							
	25							
	0이하	0	46	37	27	39	29	28
	1~4	100	53	63	51	61	63	58
	5이상	0	1	0	0	0	8	5
	0이하/4이하	0%	87%	59%	53%	64%	46%	48%
	4이하mean	1.6600	0.4848	0.4500	0.6667	0.6000	0.4348	0.3837
	SUM	100	100	100	78	100	100	91
	mean	1.66	0.54	0.45	0.6667	0.6	0.93	0.7363
Conj. (AND + Conjunction)	-1			5	4		1	1
	0	1	50	22	18	42	18	13
	1	99	35	37	27	47	53	49
	2		15	29	22	11	24	24
	3			5	5		3	3
	4			2	2		1	1
	0미만	0	0	5	4	0	1	1
	0이하	1	50	27	22	42	19	14
	SUM	100	100	100	78	100	100	91
	mean	0.99	0.65	1.13	1.1538	0.69	1.13	1.1978
(AND)	-1			5	4		2	1
	0		86	49	40	84	45	38
	1	100	14	46	34	16	51	50
	2						2	2
	0이하	0	86	54	44	84	47	39
	SUM	100	100	100	78	100	100	91
	mean	1.00	0.14	0.41	0.38	0.16	0.53	0.58
(Conjunction)	-1	1		4	4		4	4
	0	99	64	37	25	58	39	34
	1		21	46	37	31	48	46
	2		15	10	9	11	9	7
	3			3	3			
	0이하	100	64	41	29	58	43	38
	SUM	100	100	100	78	100	100	91
	mean	-0.01	0.51	0.71	0.77	0.53	0.62	0.62
Pronoun	-3			1	1			
	-2			1	1			
	-1	5		2	2		7	7
	0	95	100	87	70	100	80	72
	1			9	4		12	11
	2						1	1
	0미만(생략)	5	0	4	4	0	7	7
	1이상	0	0	9	4	0	13	12

	fail			~~23%~~	19%		~~11%~~	13%
		CLINC150 n=100						
		A	B	C	C'	D	E	E'
Length	-11							
	-10							
	-9							
	-8							
	-7							
	-6						2	2
	-5							
	-4							
	-3			1			4	4
	-2		2	5	4	1	3	3
	-1		3	2	2	4	12	10
	0		43	20	19	35	9	8
	1	31	40	42	34	48	25	22
	2	69	12	12	9	12	10	8
	3			8	6		8	8
	4			5	3		11	9
	5			3	3		6	6
	6			1			5	4
	7						2	1
	8			1	1		1	1
	9						2	1
	10							
	11							
	19							
	25							
	0이하	0	48	28	25	40	30	27
	1~4	100	52	67	52	60	54	47
	5이상	0	0	5	4	0	16	13
	0이하/4이하	0%	92%	42%	48%	67%	56%	57%
	4이하mean	1.6900	0.5700	1.0000	0.9351	0.6600	0.8452	0.7838
	SUM	100	100	100	81	100	100	87
	mean		0.57	1.24	1.1728	0.66	1.7100	1.5632
Conj. (AND + Conjunction)	-1			2	1		2	2
	0		56	30	20	50	7	1
	1	100	35	41	35	40	61	57
	2		9	26	24	10	29	27
	3			1	1		1	
	4							
	0미만	0	0	2	1	0	2	2
	0이하	0	56	32	21	50	9	3
	SUM	100	100	100	81	100	100	87
	mean	1	0.53	0.94	1.0494	0.6	1.2	1.2529
(AND)	-1							
	0		84	55	37	86	28	18
	1	100	16	45	44	14	72	69
	2							
	0이하	0	84	55	37	86	28	18
	SUM	100	100	100	81	100	100	87
	mean	1.00	0.16	0.45	0.5432	0.14	0.7200	0.7931
(Conjunction)	-1			3	2		2	2
	0	100	72	55	41	64	51	44
	1		19	37	33	26	44	40
	2		9	5	5	10	3	1
	3							
	0이하	100	72	58	43	64	53	46
	SUM	100	100	100	81	100	100	87
	mean	0.00	0.37	0.44	0.5062	0.46	0.4800	0.4598
Pronoun	-3							
	-2							
	-1			3	1		1	1
	0	100	100	94	78	100	78	74
	1			3	2		21	12
	2							
	0미만(생략)	0	0	3	1	0	1	1
	1이상	0	0	3	2	0	21	12

Opinion

- **Research Scope: focusing on "complex multi-intent" detection**
  - 당해년도 과제에서는 MID에 집중하는 것으로 의견 제안
    - Complex MI utterance의 slot-filling은 전통적인 방식으로 접근하기 어려움

e.g.

- 각 어절을 token으로 가정, token마다 slot labeling을 진행
- intent1: 라디오 작동, intent2: 볼륨 조절
  slot1: concern with intent1, slot2: concern with intent2

Simple MI setting

라디오 켜고 볼륨 줄여줘

- intent label → intent1, intent2
- slot label →

token	slot	BIO
라디오	slot1	B
켜고	slot1	I
볼륨	slot2	B
줄여줘	slot2	I

- 가상의 intent-slot labeling 가능

Complex MI setting

라디오 작게 틀어줘

- intent label → intent1, intent2
- slot label →

token	slot	BIO
라디오	slot1	B
작게	slot1	I
틀어줘	slot1, 2	?

- slot tag가 sequential 하지 않음
  (BIO tagging 불가)
- 1개의 token에 2개 이상의 label 존재

{ End Page }

HYU NLP Lab.
Hanyang University