

SBERT-Eclass: 대화 시스템의 의도 분류를 위한 개체 속성 중심의 문장 임베딩 모델

고동률⁰¹², 김창완¹, 이재윤¹, 이다희¹, 김재은^{1†}, 박상현^{2†}

(주)솔트룩스 AI Labs¹, 연세대학교 컴퓨터과학과²

{dongryul.ko, changwon.kim, jayun.lee, dahee.lee, jaeun.kim}@saltlux.com,
sanghyun@yonsei.ac.kr

SBERT-Eclass: Sentence BERT using Entity Class Information for Intent Classification in Dialogue System

Dongryul Ko, Changwon Kim, Jaeyun Lee, Dahee Lee, Jaieun Kim, Sanghyun Park
AI Labs, Saltlux Inc¹, Dept. of Computer Science, Yonsei University²

요약

의도 분류는 대화시스템에서 사용자의 발화 의도를 파악하는 중요한 과제이다. 트랜스포머 기반 언어 모델의 발전에 따라, 다양한 의도 분류 관련 연구들이 진행되고 있으며, 그 중 문장 유사도를 기반으로 사용자의 의도를 파악하는 방법이 많이 사용되고 있다. 하지만 일반적인 임베딩 모델을 활용한 문장 유사도 분석은 문장 내의 핵심어 정보에 치우친 분석을 하기 때문에, 문장 세부적인 내용보다 상위 개념의 의미론적 유사도 정보가 필요한 의도 분류에 적절하지 않다. 본 논문에서는 발화 내의 핵심어에 해당하는 개체보다 문장 전체의 의도 파악이 중요하다는 것에 착안하여, 개체 속성정보를 활용하여 의도 분류를 위한 학습 데이터셋을 구축하고, 기존 문장 임베딩 모델에 적응학습을 하는 방안을 제안하고, 이를 통해 학습한 제안 모델(SBERT-Eclass)이 기존 모델보다 우수한 성능을 보이는 것을 입증했다.

1. 서론

자연어이해 분야 중 의도 분류(Intent Classification)는 사용자 발화의 의도를 분석하는 것으로, 대화시스템에서 중요한 과제이다. 또한, 트랜스포머 기반의 언어모델이 발전함에 따라, 사용자의 의미론적(Semantic) 정보를 파악하여 의도를 분류하는 연구가 활발하게 진행되고 있다[1].

언어모델을 활용한 의도 분류 방법은 크게 분류기(Classifier)를 활용한 방법과 임베딩(Embedding) 모델을 사용하는 방법으로 나눌 수 있다. 분류기를 활용하는 방법은 언어모델에 분류기와 활성화함수를 추가하고 파인튜닝(Fine-tuning)하여, 입력 문장을 사전에 정의된 의도 클래스(Class) 중 하나로 추론하는 방법이다[2]. 분류기를 활용한 방법은 새로운 의도가 추가될 때마다 모델을 다시 학습해야 한다는 한계점이 있기 때문에, SBERT[3]와 같은 문장 임베딩 모델을 활용하여, 문장 유사도를 기반으로 의도를 추론하는 방법을 사용하는 경우가 많다[4]. 이 방법은 ‘음료주문’과 같은 의도 클래스를 설정하고, 클래스별로 ‘커피 한잔 주세요’, ‘오렌지 주스 2잔 주문하려고 합니다’와 같은 인스턴스(Instance) 문장을 구축한다. 그리고 사용자 질의와 가장 유사한 인스턴스 문장을 찾고, 인스턴스 문장의 의도 클래스를 활용하여 사용자 질의의 의도를 판단한다.

하지만, 기존 임베딩 모델은 인스턴스 문장 내 사용된 어휘의 정보를 많이 활용하기 때문에, 문장 전체의

의도를 파악하기 어렵다[5]. 예를 들어, ‘커피 한잔 주문하려고 합니다.’와 ‘오렌지 주스 2잔 주문하려고 합니다.’는 모두 ‘음료주문’이라는 동일한 의도 클래스를 가지고 있지만, 커피와 오렌지 주스라는 어휘의 차이로 인해, 유사하지 하지 않은 문장이라고 판단하여 다른 의도로 분류할 수 있다는 문제점이 있다. 따라서 인스턴스 문장 수준의 유사도 비교보다 의도 클래스 수준의 유사도 비교가 필요하다.

본 논문에서는 인스턴스 보다 상위 수준의 의미론적(High-level Semantic) 유사도인 의도 클래스 수준의 의도 분류를 위한 문장 임베딩 모델을 제안한다. SBERT의 성능을 개선한 문장 임베딩 모델인 SBERT-IQ[6]를 사용하였으며, 어휘에 집중하는 SBERT-IQ의 한계점을 개선하기 위해, 개체 속성정보를 활용한 학습 데이터셋 구축 방법과 이를 통해 학습한 SBERT-Eclass 모델을 제안한다.

2. 제안 방안

2.1. 학습 데이터셋 구축

의도 클래스 수준의 의미를 임베딩하기 위해, 개체 속성정보를 활용하여 학습 데이터셋을 구축하였다.

KLUE-NER 데이터셋[7]을 활용하여, 개체와 개체 속성정보를 같이 표기하고, 의도 클래스 수준의 유사도를 학습할 수 있는 문장 쌍을 구축하였다. KLUE-NER 데이터셋에서 개체 속성정보가 PS(인물), LC(장소), OG(기관)인 개체들을 사전(Dictionary) 구조로 저장하여,

개체를 교체(Replace)하기 위한 데이터를 구축한다. 그리고 KLUE-NER 데이터셋의 문장 중 개체명이 PS, LC, OG 중 하나로만 구성된 문장을 원본 문장으로 선정한다. 원본 문장과 동일한 개체 속성정보를 가진 문장을 임의로 추출하고, 표1과 같이 원본 문장에 임의 추출 문장의 개체를 적용하여, 개체 속성정보는 같지만 개체는 상이한 유사 문장을 구축한다. 그리고 임의 추출 문장에 원본 문장의 개체를 적용하여 동일한 개체를 사용하지만 의도는 상이한 대조 문장을 생성하였다. 학습 데이터셋은 원본-유사-대조 쌍 기준 5,134 쌍을 구축하였다.

표 1 NER 데이터를 활용한 의도분류 학습 데이터 예시

원본 문장 (Original)	"팀 조연들의 탁월한 연기와 <홍금보:PS>의 녹슬지않은 관록의 무술!"
유사 문장 (Positive)	"팀 조연들의 탁월한 연기와 <이광중:PS>의 녹슬지않은 관록의 무술!"
대조 문장 (Negative)	"넘치는 말들은 다큐멘터리를 극영화로 만들기 위한 <홍금보:PS>감독의 몸부림으로까지 보인다."

2.2. 고차원 유사도 비교를 위한 임베딩 모델 학습 방안

제안모델(SBERT-Eclass)은 구축한 학습 데이터셋을 그림 1과 같이, 기존 SBERT-IQ 모델에 트리플렛 손실함수(Triplet loss)[8]를 사용하여, 벡터 공간에서 원본 문장과 유사 문장은 가까워지고, 대조 문장은 멀어지도록 적응학습(Adaptive Learning) 하였다. 이를 통해, 인스턴스 문장이 아닌 의도 클래스 수준의 의미를 중심으로 유사도를 비교할 수 있도록 하였다.

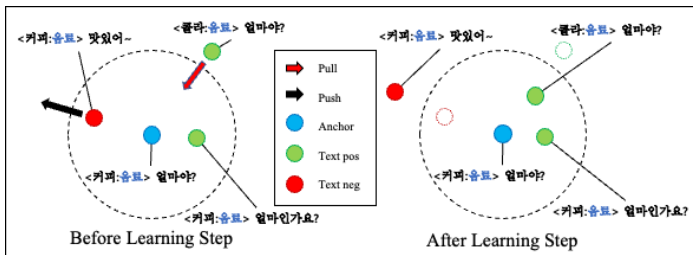


그림 1 트리플렛 손실함수를 활용한 학습 방법

3. 실험 및 평가

평가 데이터셋은 AI Hub의 한국어 대화 데이터셋을 활용하였으며, 해당 데이터셋은 도메인별 대화 데이터셋으로 발화자의 의도와 개체 속성정보가 포함되어 있다.

AI Hub 한국어 대화 데이터셋을 의도 분류 평가에 적합하게 전처리를 하였다. 첫째, 의도 클래스 별로 중복되는 인스턴스 문장을 제거하였다. 둘째, 개체 속성정보가 존재하는 발화를 인스턴스 문장으로 활용하였다. 셋째, 의도 클래스 별 인스턴스 문장이 30개 이상인 경우만 평가 대상으로 선정하였다.

이렇게 전처리를 완료한 데이터셋을 의도 클래스별

인스턴스 문장의 80%를 벡터 공간에 색인하고, 20%를 평가 데이터로 사용하였다.

표 2 평가 데이터셋 통계

도메인	전처리 전		전처리 후	
	의도	인스턴스	의도	인스턴스
음식점	356	15,726	34	2759
의류	405	15,826	40	2,967
학원	232	4,773	4	165
소매점	645	14,949	40	2,780
생활서비스	364	11,087	29	1,885
카페	133	7,859	16	1,935
숙박업	172	7,113	17	1,108
관광여가오락	160	4,949	16	707
부동산	204	8,131	27	1,507

동일한 평가 데이터를 대상으로 SBERT-IQ와 SBERT-Eclass로 인스턴스 문장을 색인하고, 코사인 유사도를 활용하여 사용자 질의와 가장 유사한 인스턴스 문장을 검색한다. 검색된 인스턴스 문장의 의도 클래스가 사용자 질의의 의도 클래스와 동일하면, 의도 분류에 성공했다고 평가했다.

평가 결과는 표3 및 그림 2과 같으며, SBERT-IQ 보다 개체 속성정보를 활용하는 SBERT-Eclass를 사용한 방법이 더 높은 성능을 보이는 것을 확인할 수 있다.

표 3 모델별 의도 분류 평가 결과

도메인	SBERT-IQ (Acc)	SBERT-Eclass (Acc)
음식점	0.596	0.701
의류	0.579	0.591
학원	0.758	0.788
소매점	0.549	0.669
생활서비스	0.645	0.721
카페	0.744	0.829
숙박업	0.739	0.725
관광여가오락	0.775	0.768
부동산	0.49	0.49

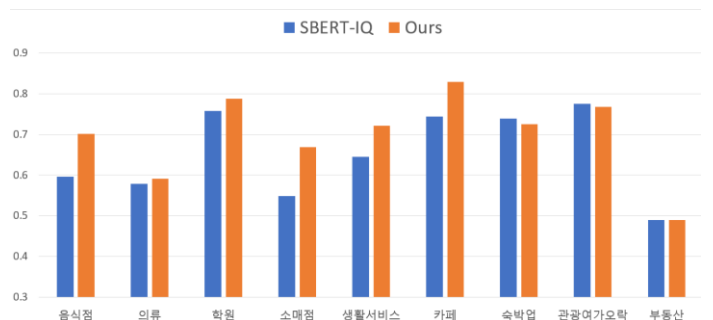


그림 2 도메인별 모델의 성능 평가 결과

표3의 숙박업과 관광여가오락 도메인에서는 SBERT-Eclass보다 SBERT-IQ의 성능이 더 높게 측정된 것을 확인할 수 있는데, 해당 데이터셋의 추론 결과를 분석한 결과 표 4와 같이 SBERT-Eclass가 적절하게 추론하였지만, 오답 처리가 된 경우를 확인할 수 있었다.

표 4 적절한 추론이지만 오답 처리된 경우 예시

도메인	분류	Input	Output	
			SBERT-IQ	SBERT-Eclass
숙박업	문장	1박에 얼마인가요?	1박에 얼마예요?	호텔 1일 숙박하는데 얼마죠?
	의도	숙박비용문의	숙박비용문의	이용요금문의
관광 여가 오락	문장	커플석으로 부탁드립니다	커플석도 있나요?	그럼 커플석으로 두 자리 주세요
	의도	커플석 문의	커플석 문의	이용 요청

추가적으로 t-SNE(t-distributed Stochastic Neighbor Embedding)를 활용하여 인스턴스 문장이 어떻게 임베딩 되는지 확인하였다. 그림 3, 4 는 카페 도메인을 대상으로 의도 클래스별 인스턴스 문장을 임베딩한 결과를 시각화한 내용이며, 그림을 통해 SBERT-IQ 보다 SBERT-Eclass 가 의도 클래스 별로 군집화가 잘 이뤄진 것을 확인할 수 있다. 따라서, 제안하는 방법이 기존 문장 임베딩 모델보다 의도 클래스 수준의 임베딩에 더 적합한 모델인 것을 확인하였다.

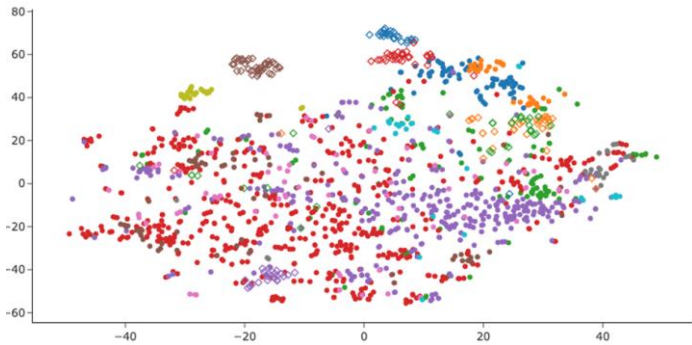


그림 3 SBERT-IQ를 활용한 임베딩 결과

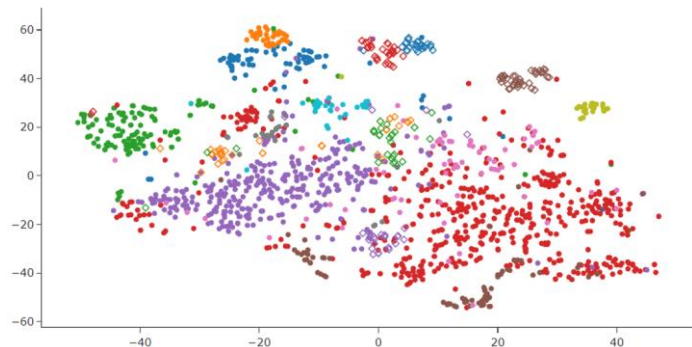


그림 4 SBERT-Eclass를 활용한 임베딩 결과

4. 결론 및 향후 연구

본 논문에서는 인스턴스 문장 수준이 아닌, 의도 클래스 수준의 의도 분류를 위해, 개체 속성정보를 활용한 학습 데이터셋 구축 방안을 제안하였다. 또한, 구축한 학습 데이터셋을 활용하여 SBERT-Eclass 모델을 학습하고, 기존 문장 임베딩 모델인 SBERT-IQ 와 성능을 비교하여, 개체 속성정보를 활용한 임베딩 모델의 의도 분류 성능의 우수함을 입증하였다.

본 논문에서 평가한 실험은 단일 도메인의 싱글턴(Single-turn) 발화를 대상으로 의도 분류 성능을 평가하였다. 향후에는 멀티 도메인, 멀티턴(Multi-turn)을 대상으로 한 의도 분류 성능향상을 위한 연구를 진행할 계획이다.

참고문헌

- [1] Samuel Louvan, Bernardo Magnini. "Recent neural methods on slot filling and intent classification for task-oriented dialogue systems: A survey," In Proceedings of the 28th International Conference on Computational Linguistics, pages 480-496, 2020.
- [2] CHEN, Yahui. "Convolutional neural network for sentence classification," Master's Thesis. University of Waterloo, 2015.
- [3] MESGAR, Mohsen, et al. "The Devil is in the Details: On Models and Training Regimes for Few-Shot Intent Classification," In Proceedings of the 29th International Conference on Computational Linguistics, pages 333-343, 2023.
- [4] Geli Fei, Bing Liu. "Breaking the Closed World Assumption in Text Classification," NAACL, 2016.
- [5] Nils Reimers, Lryna Gurevych, "Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks," In Proceedings of EMNLP-IJCNLP, 2019.
- [6] 박상민, 이재윤, 손유리, 김재은. "SBERT-IQ: 키워드 정보량을 고려한 Sentence-BERT 기반의 임베딩 모델," 한국정보과학회 학술발표논문집, pages 1058-1060, 2022.
- [7] PARK, Sungjoon, et al. Klue: Korean language understanding evaluation. NeurIPS, 2021.
- [8] Schroff, Florian, Dmitry Kalenichenko, and James Philbin, "Facenet: A unified embedding for face recognition and clustering," In Proceedings of the IEEE conference on computer vision and pattern recognition, pages 815-823, 2015.