

# 주제 임베딩을 활용한 한국어 생성 기반 챗봇

오신혁<sup>○</sup>, 김학수<sup>\*</sup>

강원대학교 컴퓨터정보통신공학과, 건국대학교 컴퓨터공학부\*  
osh7605@kangwon.ac.kr, nlpdrkim@konkuk.ac.kr

## Korean Generative Chatbot using Topic Embedding

Shinhyeok Oh<sup>○</sup>, Harksoo Kim<sup>\*</sup>

Kangwon National University, Dept. of Computer and Communication Engineering  
Konkuk University, Division of Computer Science and Engineering\*

### 요 약

챗봇은 발화에 대해 컴퓨터가 자동으로 응답하는 시스템이다. 현재 챗봇은 전체 주제에 대한 잡담(chit-chat)보다는 특정 주제에 관한 대화를 목적으로 많이 개발되고 있다. 하지만 개인이 필요로 하는 챗봇 용도에 적합한 학습 데이터는 부족하다. 이러한 상황에서 챗봇 학습을 위해 필요한 주제의 말뭉치를 대량으로 구축하는 것은 시간과 비용이 많이 소모되어 현실적으로 어렵다. 따라서 학습에 필요한 소량의 말뭉치만 사용하더라도 주제에 적합한 응답을 할 수 있는 챗봇이 필요하다. 이에 본 논문은 챗봇의 목적과 관련 없는 대량의 말뭉치와 소량의 주제 기반 말뭉치를 이용하여 높은 성능을 끌어낼 수 있는 주제 임베딩 방법을 제안한다.

**주제어:** 챗봇, 주제 기반 챗봇, 주제 임베딩

### 1. 서론

챗봇은 특정 발화에 대해 컴퓨터가 자동으로 응답하는 시스템을 말한다. 간단한 질의응답 업무에서 복잡한 상담 관련 업무에 이르기까지 이전보다 다양한 분야에서 챗봇을 도입하고 있으며 이에 따른 챗봇 연구의 필요성은 꾸준히 증가하고 있다. 현재까지 사용되고 있는 챗봇은 특정 목적을 갖고 서비스하는 것이 일반적이다. 그러나 생성 기반 챗봇 연구는 주로 특정 주제의 학습 데이터만을 이용하거나 주제와 관계없이 다양한 학습 데이터를 이용하여 연구가 진행되고 있다. 전자의 경우 실제 챗봇을 서비스하기 위해서는 주제에 알맞은 학습 데이터가 많지 않기 때문에 필요한 챗봇 주제에 대한 데이터를 구축하기 위해 시간과 비용이 많이 소모된다. 후자는 전자보다 비교적 많은 데이터를 쉽게 활용할 수 있지만, 주제와 관련 없는 데이터도 모두 사용하기 때문에 챗봇 모델이 대화 주제에 적합하지 않은 응답을 출력할 가능성이 크다. 이에 따라 서비스할 챗봇의 주제와 관련된 소량의 말뭉치와 주제와 관련 없는 대량의 말뭉치를 이용해 효과적인 성능을 낼 수 있는 주제 기반 챗봇 연구가 필요하다. 본 논문은 주제 기반 챗봇에서 주제와 관련 없는 일반적인 대량의 말뭉치와 소량의 주제 기반 말뭉치를 이용하여 높은 성능을 끌어낼 수 있는 주제 임베딩 방법을 제안한다.

### 2. 관련 연구

과거부터 생성 기반 챗봇의 연구는 시퀀스-투-시퀀스(sequence-to-sequence) 모델로 많은 연구가 수행되었다 [1-3]. 하지만 시퀀스-투-시퀀스 기반 모델은 짧고 안전한 응답만을 생성하거나, 의미상 적합하지 않은 응답을 하는 문제점이 종종 보고되고 있다 [4,5]. 이러한 문제점을 해결하기 위해 VAE [6]와 WGAN [7] 기반의 챗봇 모델과 같이 잠재 공간을 학습하는 다양한 챗봇 모델 연구가 지속해서 이루어지고 있다 [8,9]. 그러나 VAE 기반 모델은 디코더가 학습된 잠재 공간을 반영하지 않거나 표준 정규 분포로 잠재 공간을 단순화하는 경향을 띤다. 따라서 본 논문에서는 WGAN 기반의 챗봇 모델 [9]을 제안 방법 실험에 사용했다. 또한, 한국어 챗봇에 특성을 반영할 수 있는 연구도 꾸준히 진행되고 있다. 챗봇에 감정을 반영하고 사용자의 감정에 적합한 응답을 하는 감정 엔진에 관한 연구 [10]와 개인 특성을 의미하는 페르소나를 챗봇에 반영하는 연구 [11-13] 등 다양한 연구가 수행되고 있다. 하지만 대화의 가장 기초가 되는 요소인 주제 반영에 관한 연구는 부족하다. 하나의 챗봇이 서비스하는 분야는 한정적이기 때문에 전체 주제에 대해 응답할 수 있는 챗봇보다는 목표한 주제에 적합한 응답을 할 수 있는 챗봇이 필요하다. 따라서 본 논문은 생성 기반 챗봇에 주제를 반영하게 하기 위한 주제 임베딩 방법을 제안하고 실험을 통해 효과를 입증한다.

### 3. 주제 임베딩을 활용한 한국어 생성 기반 챗봇

표 1 학습 데이터 통계

(단위: 개)

일반	음식점	의류	학원	소매점	생활 서비스	카페	숙박업	관광여가 오락
20,008	5,469	4,779	1,268	4,990	3,134	2,536	2,324	1,721

### 3.1. 주제 임베딩 생성 방법

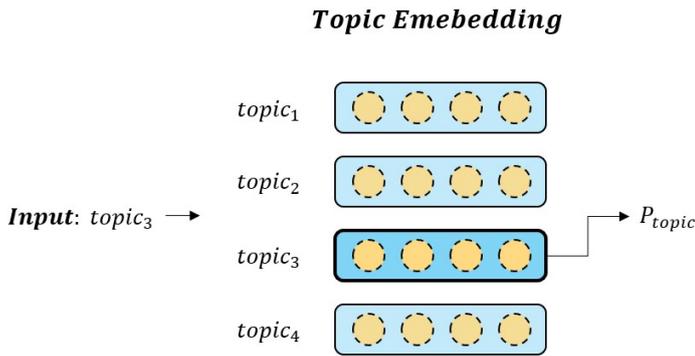


그림 1 주제 임베딩 방법

그림 1은 본 논문에서 제안한 주제 임베딩 방법이다. 현재 입력된 담화에 대한 주제를 임베딩 하여  $P_{topic}$ 을 생성한다. 생성한  $P_{topic}$ 은 생성 기반 챗봇 모델의 디코딩(decoding) 단계에서 사용한다.

### 3.2. 주제 임베딩 디코더 반영 방법

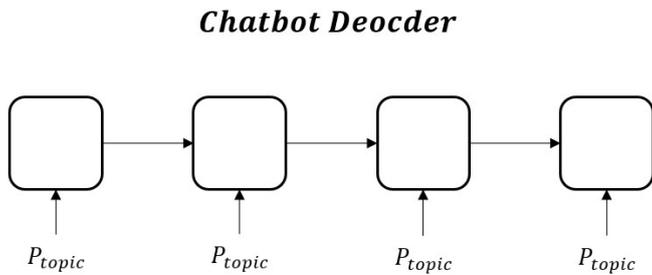


그림 2 주제 임베딩 디코더 반영 방법

주제 임베딩 방법을 통해 생성된  $P_{topic}$ 은 생성 기반 챗봇의 디코더의 매 시간 단계(time-step)마다 반영되어 디코더에서 생성하는 토큰들이 담화 주제와 같은 경향을 띠게 한다. 또한, 역전파(backpropagation) 과정을 통해 주제 임베딩( $P_{topic}$ )이 단어 임베딩과 생성 기반 챗봇 모델의 다른 매개변수(parameter)들과 함께 학습될 수 있게 한다.

## 4. 실험 및 평가

### 4.1. 실험 데이터

본 논문의 실험을 위해 국립국어원에서 공개한 구어 말뭉치와 한국정보화진흥원에서 공개한 한국어 대화 말뭉치를 사용했다. 구어 말뭉치는 ‘일반’ 주제의 학습 데이터로 사용했으며 그 외 주제는 한국어 대화 말뭉치를 이용했다. 학습에 사용한 발화 수(turn)는 4개로 제한했으며, 이에 따른 데이터 통계는 표 1과 같다.

### 4.2. 실험 방법

본 논문에서 제안한 주제 임베딩 방법의 실험을 위해 [13]의 챗봇 모델을 사용했다. 해당 모델에 주제 임베딩 방법을 적용하여 정량평가와 정성평가를 수행했다. ‘단일 주제’ 모델은 표 1의 학습 데이터 중 음식점 주제 데이터만 이용하여 학습한 모델이고, ‘다중 주제’ 모델은 표 1의 학습 데이터 전체를 이용해 학습한 모델이다. ‘다중 주제’에서 주제 임베딩을 적용한 모델은 본 논문의 제안 모델로 학습 데이터를 표 1에 기재된 주제별로 제안 방법을 적용하여 학습한 모델이다.

### 4.3. 정량평가 결과

본 논문에서는 주제 임베딩 방법의 검증에 위해 표 1의 음식점 주제 데이터 608개를 정량평가 데이터로 사용했고 이에 따른 정량평가 결과를 표 2에 기재했다. 정량평가 척도로 BLEU[14], Embedding Score[15], Distinct[4], 응답 토큰 수를 사용했다. BLEU는 모델이 생성한 응답과 정답 간의 n-gram 중첩량을 표현해주는 수치다. BLEU 척도를 계산할 때는 스무딩(smoothing)기법[16]을 적용하여  $n < 3$ 에 대한 BLEU를 계산한다. Embedding Score는 응답과 정답 임베딩 간의 코사인 유사도를 나타낸다. Distinct는 응답의 다양성을 계산하는 척도로 사용된다. Intra-dist는 하나의 응답 내에서 고유한 n-gram 비율의 평균을 의미하며, Inter-dist는 모든 응답 내에서 고유한 n-gram의 비율을 의미한다.

표 2에서 다중 주제의 데이터를 주제 임베딩을 사용하여 학습한 경우 주제 임베딩을 사용하지 않았을 때보다 다양성을 의미하는 Distinct를 제외한 모든 척도에서 우수한 성능을 보였다. 주제 임베딩을 사용하면 디코더가 응답을 생성할 때 현재 주제와 같은 경향을 띠게 하기 때문에 주제 임베딩을 제외하고 학습할 경우보다 생성되는 토큰의 다양성은 저하되지만 응답 자체의 품질과 관련 있는 척도인 BLEU 점수와 Embedding 점수, 생성된 토큰의 수에 대해서는 우수한 성능을 보였다. 또한, 단일 주제만을 학습할 때보다 다중 주제의 데이터에서 주제 임베딩을 사용했을 경우 BLEU 점수와 생성되는 토큰의 수에 대해서 유의미한 차이를 보였으며 Embedding 점수

표 2 정량평가 결과

평가 기준 학습 방법	BLEU	Embedding			Distinct				응답 토큰 수
	Smoothing	Avg	Greedy	Extrema	Intra-dist		Inter-dist		
					Unigram	Bigram	Unigram	Bigram	
단일 주제	0.180	<b>0.397</b>	0.413	<b>0.247</b>	<b>0.926</b>	<b>0.991</b>	0.099	0.282	5.92
다중 주제	0.149	0.339	0.382	0.207	0.921	0.986	<b>0.127</b>	<b>0.351</b>	6.10
다중 주제 (+주제 임베딩)	<b>0.203</b>	0.394	<b>0.417</b>	0.219	0.901	0.984	0.103	0.295	<b>7.56</b>

와 다양성을 의미하는 Distinct 점수에서는 유사한 성능을 보였다.

#### 4.4. 정성평가 결과

표 3 정성평가 결과

평가 기준 학습 방법	문법 (0/1)	의미 (0/0.5/1)	주제 적합성 (0/0.5/1)
단일 주제	0.965	0.704	<b>0.675</b>
다중 주제	0.946	0.578	0.489
다중 주제 (+주제 임베딩)	<b>0.971</b>	<b>0.737</b>	0.667

생성 기반 챗봇 모델은 정량평가만으로 모델의 성능을 측정하기는 어렵다. 따라서 본 논문에서는 주제 임베딩 방법의 검증에 위해 정성평가를 추가로 진행했다. 평가에는 학습에 사용하지 않은 음식점 주제 데이터 62문장을 사용했다. 평가자는 총 3명이며 3인의 점수를 평균하여 표 3에 기재했다. 평가 기준 중 문법 점수는 생성된 응답의 문법이 적합하지 않으면 0점, 적합하면 1점으로 평가했으며 의미 점수는 입력 담화에 의미상으로 적합한 응답인지에 대해 0, 0.5, 1점으로 평가했다. 주제 적합성은 응답이 현재 선택한 주제인 음식점에 적합한 응답인지에 대해 0, 0.5, 1점으로 평가했다. 주제 임베딩을 사용한 모델이 비교 모델들보다 문법, 의미상으로 우수함을 보였고, 주제 적합성은 단일 주제만을 학습한 모델의 성능과 유사함을 보였다.

#### 4.5. 응답 예제

챗봇 응답 예제는 부록에 별도로 첨부했다. 챗봇 모델로 사용한 이중 WGAN 챗봇[13]과 같이 복합 표현 단위 단어 임베딩[17]을 사용했기 때문에 미등록어나 숫자로 표현되는 부분은 토큰 앞에 ###을 부착했고 생성 시 음절로 생성한 부분은 토큰 앞에 #을 부착했다. 예제를 보면 주제 임베딩을 사용한 모델이 비교 모델들보다 전반적으로 상세하고 올바른 응답을 하는 것을 보였다.

#### 5. 결론 및 향후 연구

본 논문은 생성 기반 챗봇에서 사용할 수 있는 주제 임베딩 방법을 제안했고 실험을 통해 주제 기반 챗봇에서 여러 주제의 다량의 말뭉치와 챗봇 주제에 관련된 소량의 말뭉치를 이용해 제안 모델의 구조로 학습하면 효과적임을 보였다. 또한, 주제 임베딩 방법을 사용하면 다양한 주제의 챗봇을 하나의 모델로 사용하거나 하나의 주제에 대한 챗봇을 소주제로 분리하여 사용하는 것도 가능하므로 이와 같은 연구도 추후 진행할 예정이다.

#### 참고문헌

- [1] 김시형, 김학수, “의사 형태소 단위 채팅 시스템”, 제 28회 한글 및 한국어 정보처리 학술대회 논문집, pp.263-267, 2016.
- [2] 김진태, 김시형, 김학수, 이연수, 최맹식, “색인어 인코딩과 음절 디코딩에 기반한 생성 채팅 모델”, 제 29회 한글 및 한국어 정보처리 학술대회 논문집, pp.125-129, 2017
- [3] 김진태, 김학수, 최맹식, 이연수, 권오욱, 김영길, “비지도 사전 학습을 이용한 한국어 채팅 시스템”, 2018 한국컴퓨터종합학술대회 논문집, pp.631-633, 2018.
- [4] J. Li, M. Galley, C. Brockett, J. Gao, B. Dolan, “A Diversity-Promoting Objective Function for Neural Conversation Models”, Proc. of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human

- Language Technologies, pp. 110-119, 2016.
- [5] S. Sato, N. Yoshinaga, M. Toyoda, M. Kitsuregawa, "Modeling Situations in Neural Chatbots", Proc. of the 55th Annual Meeting of the Association for Computational Linguistics, Student Research Workshop, pp. 120-127, 2017.
- [6] D. P. Kingma, M. Welling, "Auto-encoding variational bayes", 2nd International Conference on Learning Representations, Conference Track Proceedings, 2014.
- [7] M. Arjovsky, S. Chintala, L. Bottou, "Wasserstein Generative Adversarial Networks", Proceedings of the 34th International Conference on Machine Learning, Vol. 70, pp. 214-223, 2017.
- [8] X. Gu, K. Cho, J.W. Ha, S. Kim, "DialogWAE: Multimodal Response Generation with Conditional Wasserstein Auto-Encoder", 7th International Conference on Learning Representations, 2018.
- [9] J. Kim, S. Oh, O.-W. Kwon, H. Kim, "Multi-Turn Chatbot Based on Query-Context Attentions and Dual Wasserstein Generative Adversarial Networks", Appl. Sci., 3908, 2019.
- [10] 이성민, "감정 모델 기반 챗봇 구현 연구", 석사 학위논문, 숭실대학교, 2018.
- [11] J. Li, M. Galley, C. Brockett, "A Persona-Based Neural Conversation Model", Proc. of the 54th Annual Meeting of the Association for Computational Linguistics, Vol. 1, pp. 994-1003, Aug. 2016.
- [12] 오신혁, 김진태, 박영민, 김선아, 이정엄, 김학수, "생성 기반 챗봇 시스템을 위한 페르소나 반영 방법", 2019 한국컴퓨터종합학술대회, pp. 1761-1763, 2019.
- [13] 오신혁, 김진태, 김학수, 이정엄, 김선아, 박영민, 노명호, "Dual WGAN 기반 페르소나 Multi-Turn 챗봇", 제 31회 한글 및 한국어 정보처리 학술대회 논문집, pp. 49-53, 2019.
- [14] K. Papineni, S. Roukos, T. Ward, and W. J. Zhu, "BLEU: a method for automatic evaluation of machine translation", Proc. of the 40th annual meeting on association for computational linguistics, pp. 311-318, 2002.
- [15] C. W. Liu, R. Lowe, I. V. Serban, M. Noseworthy, L. Charlin, and J. Pineau, "How NOT To evaluate your dialogue system: An empirical study of unsupervised evaluation metrics for dialogue response generation", Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing, pp. 2122-2132, 2016.
- [16] B. Chen, C. Cherry, "A systematic comparison of smoothing techniques for sentencelevel BLEU", Proc. of the Ninth Workshop on Statistical Machine Translation, pp. 362-367, 2014.
- [17] 김진태, 이현구, 김학수, "소량의 대화 말뭉치에서 학습 가능한 효과적인 생성 기반 챗봇 모델", 정보과학회논문지, 46(3), pp.246-252, 2019.

부록

표 4 응답 예제 1

이전 대화		치킨 <SP> 시키 려고 <SP> 하 는데요
		네 <SP> 어떤 것 로 <SP> 하 아 드리 르까요 ?
		###후라이드 <SP> 반 <SP> 양념 <SP> 반 으로 <SP> 하 아 주 시 어요
답변	단일 주제	네 <SP> 알 겠 습니다
	다중 주제	네
	다중 주제 (+주제 임베딩)	네 <SP> 더 <SP> 필요 하 나 <SP> 것 은 <SP> 없 으시 어요 ?

표 5 응답 예제 2

이전 대화		그럼 <SP> 내일 <SP> 오후 <SP> ###2 시 에 <SP> 예약 <SP> ###부탁드릴게요
		네 <SP> 몇 <SP> 분 으로 <SP> 예약 <SP> 돕 아 드리 르까요 ?
		###5 명 이 <SP> 방문 하 르 <SP> 예정 이 ㅂ니다
답변	단일 주제	#3 #0 분 <SP> 정도 <SP> 기다리 시 어야 <SP> 되 ㅂ니다
	다중 주제	네 , <SP> 할인 <SP> 되 시 어요
	다중 주제 (+주제 임베딩)	네 <SP> 단체 로 <SP> 안내 하 아 <SP> 드리 겠 습니다

표 6 응답 예제 3

이전 대화		비빔밥 <SP> 종류 는 <SP> 없 어요 ?
		비빔밥 <SP> 있 어요
		볶음밥 은 <SP> 있 어요 ?
답변	단일 주제	없 습니다
	다중 주제	없 어요
	다중 주제 (+주제 임베딩)	볶음밥 은 <SP> 따로 <SP> 없 어요

표 7 응답 예제 4

이전 대화		네 <SP> ###12 시 <SP> 정각 에 <SP> 받 으러 오 르게요
		성함 하 고 <SP> 전화번호 <SP> 적 어 주 시 어요
		계산 은 <SP> 카드 로 <SP> 하 르게요
답변	단일 주제	네
	다중 주제	네 <SP> 감사 하 ㅂ니다
	다중 주제 (+주제 임베딩)	네 <SP> #1 #5 #0 #0 #0 원 이 ㅂ니다

표 8 응답 예제 5

이전 대화		###4 명 <SP> 자리 <SP> 있 습니까 ?
		네 <SP> 룸 이 <SP> 아니 어도 <SP> ###괜찮으신가요?
		네 <SP> 괜찮 습니다
답변	단일 주제	포장 이 시 르가요 ?
	다중 주제	네
	다중 주제 (+주제 임베딩)	네 <SP> 잠시 만 <SP> 기다리 어 <SP> 주 시 어요