

# CNN-LSTM 신경망을 이용한 발화 분석 모델

김민경<sup>o</sup>, 김학수

강원대학교 컴퓨터정보통신공학과

kmink0817@kangwon.ac.kr, nlpdrkim@kangwon.ac.kr

## Utterance Intention Analysis Using CNN-LSTM Neural Network

Min-Kyoung Kim<sup>o</sup>, Harksoo Kim

Kangwon National University Computer and Communication Engineering

### 요 약

대화시스템이 적절한 응답을 제시해 주기 위해서는 사용자의 의도를 분석하는 것은 중요한 일이다. 사용자의 의도는 도메인에 독립적인 화행과 도메인에 종속적인 서술자의 쌍으로 나타낼 수 있다. 사용자 의도를 정확하게 분석하기 위해서는 화행과 서술자를 동시에 분석하고 대화의 문맥을 고려해야 한다. 본 논문에서 제안하는 모델은 합성곱 신경망에서 공유 계층을 이용하여 화행과 서술자간 상호작용이 반영된 발화 임베딩 모델을 학습한다. 그리고 순환 신경망을 통해 대화의 문맥을 반영하여 발화를 분석한다. 실험 결과 제안 모델이 이전 모델들 보다 높은 성능 (F1-measure로 화행에 대해 0.973, 서술자 0.919)을 보였다.

주제어: 화행, 서술자, 공유계층, CNN-LSTM

### 1. 서론

목적 지향 대화시스템은 한정된 도메인 안에서 사용자 발화(utterance)에 대해 적절한 응답을 제시해 주는 시스템을 말한다. 목적 지향 대화시스템은 사용자와 자연스럽게 의사소통하기 위해 발화에 내포된 화자의 의도를 분석하는 것이 중요하다. 본 논문에서는 발화에 내포된 의도를 화행(speech-act)과 서술자(predicator)의 쌍으로 표현한다. 화행은 도메인에 독립적으로 사용자가 전달하고자 하는 일반적인 의도를 나타낸다. 서술자는 도메인에 종속적이며 주된 서술어의 의미 범주를 나타낸다. 표 1은 일정 관리 도메인에서 목적 지향 발화의 예와 해당 의도를 보여준다.

표 1 목적 지향 대화의 예

	발화	의도	
		화행	서술자
U	(1) 안녕~	Greeting	Null
S	(2) 무엇을 도와드릴까요?	Opening	Null
U	(3) 약속 잡아줘	Request	Update -appointment
S	(4) 날짜는 언제로 할까요?	Ask-ref	Update-date
U	(5) 10월 8일	Response	Update-date

화행과 서술자는 문맥에 의존적이기 때문에 하나의 발화만으로 추론하는 것은 매우 어렵다. 예를 들어 표 1의 발화 (5)는 두 가지 의도로 분석이 가능한데 현재 설정되어 있는 일정을 알려주는 “Inform & Select-date”와 일정이 뭐로 변경되었는지를 묻는 질문에 대해 답해주는 “Response & Update-date”가 될 수 있다. 이러한 모호성을 해결하기 위해 발화 (5)에 문맥이 반영되어야 한다. 위 예시에서 바로 이전 발화인 (4)를 고려하면 발화 (5)의 올바른 의도인 “Response & Update-date”를 선택할 수 있다.

### 2. 관련 연구

사용자 의도를 분석하기 위해 다양한 자질에 기반을 둔 기계 학습 모델들이 제안되었지만, 기존의 연구들은 주로 발화의 화행 분류만을 다루었거나[1][2] 화행과 서술자를 개별적으로 다루어왔다[3]. 그러나 사용자의 의도를 더 정확히 파악하기 위해 화행과 서술자를 동시에 식별할 필요가 있다. 관련 연구로 [4]는 서술자 예측 결과가 화행 분류를 위한 입력으로 사용되는 통합 신경망 모델을 제안하였다. [5]는 통합 모델의 성능 향상을 위해 상호 재학습 방법을 제안하였다. [5]가 제안한 모델은 화행 분류 모델과 서술자 분류 모델로 나뉘며 학습 동안 한 모델의 출력 값을 다른 모델의 입력 자질로 사용한다. 본 논문에서는 합성곱 신경망(Convolutional Neural Network)[6]와 LSTM 순환 신경망(Long Short-Term Memory Recurrent Neural Network)[7]을 이용하여 화행과 서술자를 동시에 분석하는 발화 분석 모델을 제안한다. 제안된 모델은 합성곱 신경망을 기반으로 한 새로운 발화 임베딩 방법을 이용하여 화행과 서술자간의 상호작용이 가능하게 한다. 그리고 LSTM 순환 신경망을 기반으로 대화의 문맥을 반영하여 의도 분석의 성능을 향상시킨다.

### 3. 발화 분석 모델

본 논문에서 제안하는 발화 분석 모델의 구조도는 [그림 1]과 같다. 발화 분석 모델은 화행 분류 모델(SA Classifier)과 서술자 분류 모델(PR Classifier), 발화 임베딩 모델(Utterance embedding model)로 구성된다.

대화의 문맥을 고려하여 화행과 서술자를 분류하기 위해 각 분류 모델은 LSTM 순환 신경망을 적용한다. 발화 임베딩 모델을 이용하여  $m$ 개의 발화( $Utterance_{1,m}$ )에 대해 화행 분류를 위한 임베딩 벡터( $Emb_S^{1,m}$ )와 서술자 분류를 위한 임베딩 벡터( $Emb_P^{1,m}$ )를 얻는다. 최종적으로 화행 분

류 모델과 서술자 분류 모델은 각각의 임베딩 벡터를 입력받아 화행과 서술자를 출력한다.

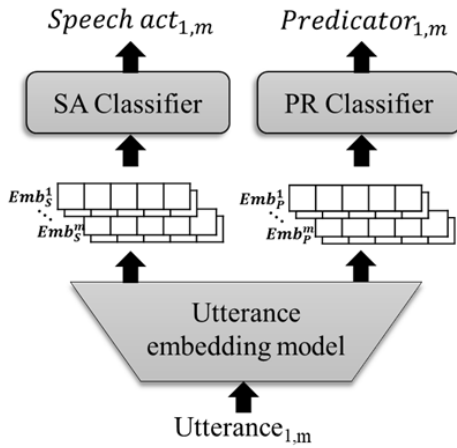


그림 1. 제안 모델의 구조도

아래 [그림 2]는 발화 임베딩을 위한 합성곱 신경망의 구조이다.

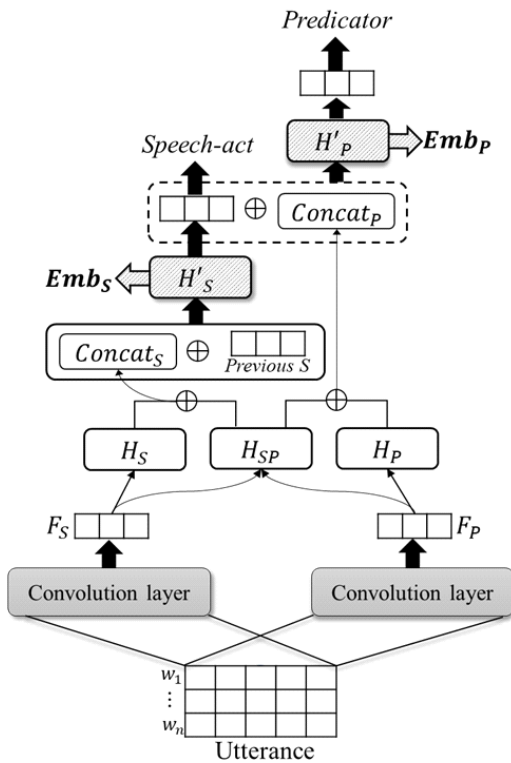


그림 2. 합성곱 신경망을 이용한 발화 임베딩 모델

[그림 2]에서 입력 발화의 각 단어  $w_n$ 은 50차원의 Word2Vec 임베딩 벡터이다[8]. 입력된 발화는 두 개의 독립된 convolution 계층을 통해 화행과 서술자에 적합한 자질 벡터  $F_S, F_P$ 를 생성한다. 은닉 노드  $H_X$ 는 자질 벡터  $F_X$ 만을 입력으로 하며( $X \in S, P$ ),  $H_{SP}$ 는  $F_S, F_P$ 를 모두 입력으로 한다. 이는 공유 계층으로 화행과 서술자의

조합된 정보를 추상화 할 수 있다.  $F_X$ 를 입력으로 하는 은닉 노드들은  $H'_X$ 의 입력이 된다. 예를 들어, [그림 1]에서 은닉 노드  $H_S$ 와  $H_{SP}$ 가  $H'_S$ 의 입력이 된다. 화행을 분류할 때, 입력된 이전 화행(Previous S)을 자질로 사용하며 서술자를 분류할 때는 모델이 예측한 현재 화행을 자질로 사용한다. 모델을 학습할 때 예측 화행과 정답 화행간의 오류가 화행과 관련된 노드들(i.e.,  $H'_S, H_S, H_{SP}$ )로 부분적 역 전파되며, 같은 방식으로 서술자에 대한 오류가 역 전파된다. 학습이 완료된 임베딩 모델에서  $H'_S$ 와  $H'_P$ 를 각각 화행과 서술자를 분류하기 위한 임베딩 값  $Emb_S, Emb_P$ 로 이용한다.

## 4. 실험 및 평가

### 4.1 실험 준비

본 논문에서는 실험을 위해 일정 관리 도메인 대화 말뭉치를 실험 데이터로 사용하였다[9]. 실험 데이터는 899개의 대화로 구성되어 있으며 전체 발화 수는 10,043개(대화 당 평균 11개의 발화)이다. 실험에서 화행과 서술자 범주는 각각 11개, 47개를 사용하였다. 실험은 10배 교차 검증을 시행하였다. 제안 모델의 성능을 평가하기 위한 척도로 정확도(accuracy), 각 클래스별 정확률의 평균(macro precision:MP), 각 클래스별 재현율의 평균(macro recall:MR), macro F1-measure를 사용하였다.

제안 모델에 세부적인 파라미터는 아래 표 2와 같다.

표 2 모델 파라미터

	발화 임베딩 모델	발화 분석 모델
num epoch	300	300
batch size	64	16
learning rate	0.0001	0.0001

### 4.2 실험 결과

제안 모델의 성능을 확인하기 위해 동일한 실험 데이터를 사용한 다른 모델들과 비교하였다. 표 3은 제안 모델과 비교 모델들 사이의 성능 차이를 보여준다.

표 3 제안 모델과 이전 모델들의 성능 비교

		accuracy	MP	MR	F1
화행	제안모델	<b>0.985</b>	<b>0.989</b>	<b>0.958</b>	<b>0.973</b>
	[5]	0.941	0.878	0.915	0.896
	[4]	0.861	-	-	-
서술자	제안모델	<b>0.975</b>	<b>0.936</b>	<b>0.903</b>	<b>0.919</b>
	[5]	0.909	0.827	0.768	0.796
	[4]	0.738	-	-	-

[5]는 SVM을 기반으로 화행 분류모델과 서술자 분류모델의 출력이 서로의 입력이 되어 학습하는 상호 재학습 방법을 이용한 모델의 성능이고 [4]는 서술자 예측 결과가 화행 분류를 위한 입력이 되는 통합 신경망 모델의

성능 결과이다. 표 3에서 제안 모델은 자질 추출 및 선택에 많은 비용을 들이지 않고도 뛰어난 성능을 보이는 것을 확인할 수 있다. 특히 상호 재학습 방법을 이용하지 않고도 [5]보다 높은 성능을 보였다. 이는 공유계층을 이용한 임베딩한 방법이 화행과 서술자가 서로 상호작용하는데 도움을 준다는 것을 보여준다.

## 5. 결론

본 논문에서는 발화를 분석하기 위해 합성곱 신경망과 LSTM 순환 신경망을 결합한 모델을 제안하였다. 합성곱 신경망을 통해 화행과 서술자간의 상호작용이 반영되게 발화를 임베딩하고 LSTM 순환 신경망을 이용하여 대화의 문맥을 반영하였다. 실험결과 제안 모델이 자질 튜닝 없이도 비교 모델들 보다 뛰어난 성능을 보였다.

## 감사의 글

이 논문은 2016년도 정부(미래창조과학부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임 (No.2016R1A2B4007732)

## 참고문헌

- [1] 은종민, 이성욱, 서정연, “지지벡터기계 (Support Vector Machines)를 이용한 한국어 화행 분석”, 정보과학회논문지(B), 제12권, 제3호, pp.365-368, 2005.
- [2] S. Kang, H. Kim, J. Seo, “A reliable Multidomain model for speech act classification”, Pattern Recognition Letters, vol.31(1), pp.71-74, 2010.
- [3] H. Lee, H. Kim, J. Seo, “Domain Action Classification Using a Maximum Entropy Model in a Schedule Management Domain”, AI Communications, vol.21(4), pp.221-229, 2008.
- [4] H. Lee, H. Kim, J. Seo, “An Integrated Neural Network Model for Domain Action Determination in Goal-oriented Dialogues”, JIPS, vol.9(2), pp.259-270, 2013.
- [5] C. Seon, H. Kim, J. Seo, “Improving Domain Action Classification in Goal-oriented Dialogues Using a Mutual Retraining Method”, Pattern Recognition Letters, vol.45, pp.154-160, 2014
- [6] A. Krizhevsky, I. Sutskever, GE. Hinton, “Imagenet classification with deep convolutional neural networks”, Advances in neural information processing systems, pp.1097-1105, 2012.
- [7] S. Hochreiter, J. Sschmidhuber, “Long short-term memory”, Neural computation, 1997.
- [8] Y. Goldberg, O. Levy, “word2vec Explained: deriving Mikolov et al.’s negative-sampling word-embedding method”, arXiv preprint

arXiv:1402.3722, 2014.

- [9] H. Kim, C. Seon, J. Seo, “Review of Korean Speech Act Classification: Machine Learning Methods”, Computing Science and Engineering, vol.5(4), pp.288-293, 2011