

사전학습 모델을 이용한 음식업종 고객 발화 의도 분류 분석

김준희^o, 임희석^{*}

^o고려대학교 인공지능융합학과,

^{*}고려대학교 컴퓨터정보통신대학원

e-mail: booguk300@korea.ac.kr^o, limhseok@korea.ac.kr^{*}

Analysis of utterance intent classification of customer in the food industry using Pretrained Model

Jun Hoe Kim^o, HeuiSeok Lim^{*}

^oDept. of Applied Artificial Intelligence, Korea University,

^{*}Graduate School of Computer & Information Technology, Korea University

● 요약 ●

기존 자연어 처리 모델은 문맥 단위 단어 임베딩을 처리하지 못하는 한계점을 가지고 있는 한편 최근 BERT 기반 사전학습 모델들은 문장 단위 임베딩이 가능하고 사전학습을 통해 학습 효율이 비약적으로 개선되었다는 특징이 있다. 본 논문에서는 사전학습 언어 모델들을 이용하여 음식점, 배달전문점 등 음식 업종에서 발생한 고객 발화 의도를 분류하고 모델별 성능을 비교하여 최적의 모델을 제안하고자 한다. 연구결과, 사전학습 모델의 한국어 코퍼스와 Vocab 사이즈가 클수록 고객의 발화 의도를 잘 예측하였다. 한편, 본 연구에서 발화자의 의도를 크게 문의와 요청으로 구분하여 진행하였는데, 문의와 요청의 큰 차이점인 ‘물음표’를 제거한 후 성능을 비교해본 결과, 물음표가 존재할 때 발화자 의도 예측에 좋은 성능을 보였다. 이를 통해 음식 업종에서 발화자의 의도를 예측하는 시스템을 개발하고 챗봇 시스템 등에 활용한다면, 발화자의 의도에 적합한 서비스를 정확하게 적시에 제공할 수 있을 것으로 기대한다.

키워드: 의도 분류(intent classification), 사전학습(pretrained), 단어집 크기(vocab size)

I. Introduction

최근 다양한 분야에서 딥러닝 기반 연구들이 진행되고 있으며, 자연어 처리 분야도 딥러닝 기반으로 세부 분야로 나뉘어 연구가 진행되고 있다. 세부 분야 중 하나인 발화자 의도 분류에서도 Bi-LSTM, RCNN 등의 딥러닝 모델에 연구[1]가 진행되었지만, 사전학습 모델 기반 연구는 부족한 실정이다. 본 논문에서는 사전학습 모델 중 BERT[2], BERT 기반으로 한국어 코퍼스를 학습한 KoBERT, ELECTRA[3] 기반으로 한국어 코퍼스를 학습한 KoELECTRA의 발화 의도 분류 성능을 비교한다. 사용한 데이터는 한국과학기술정보연구원 음식업종 소상공인 대화 데이터이며, 총 15,726개 문장 중 고객의 질문 발화만 추출하여 7,434 문장을 이용하였다.

II. Preliminaries

1. Related works

1.1 국내외 동향

딥러닝 모델 이전에는 SVM, HMM 모델 등을 이용하여 이전 발화의 의도를 고려하는 것이 발화 의도 예측 모델의 성능을 향상시킨다는 것을 보였다.[4] 딥러닝 모델 연구에서는 Bi-LSTM과 RCNN을 이용하여 단일 발화가 아닌 문맥을 고려한 발화 의도 예측 모델이 딥러닝 모델 이전보다 좋은 성능을 보여준다.[1]

III. The Proposed Scheme

1. Research Environment

1.1 데이터 수집

본 연구에서는 대화 문맥이 아닌 한 문장 단위로만 모델 성능을 비교하였으며, 데이터는 음식점종 중 배달음식점, 홀서빙음식점, 셀프 서비스 3개로 구성되어 있다. 총 15,726개 문장 중 고객의 질문 발화만 추출하여 7,434 문장으로 연구를 진행하였다.

1.2 데이터 처리

기본 데이터의 경우 유사한 클래스가 존재함에도 불구하고 총 352개의 클래스로 분류되어 있고, 일부 클래스에 편향되어 있어 1차적으로는 중앙값인 6개 이하 label에 대해서는 기타로 처리하여 159개로 통합하였다. 이후 2차 수작업을 통해 비슷한 유형의 클래스를 큰 카테고리 분류하여 19개의 클래스로 분류하여 모델 성능을 비교하였다. 추가적인 연구로는 발화자의 의도인 문의와 요청이 ‘물음표’라는 의문문의 문장부호에 의해 구분이 되는지 확인하기 위해 물음표를 제거한 후 KoELECTRA 모델 성능을 비교했다.

2. Analysis Method

데이터셋 중 70%는 훈련 데이터, 20%는 평가 데이터, 10%는 테스트 데이터로 분리하여 연구를 진행하였다. 모델은 BERT, KoBERT, KoELECTRA 총 3개를 활용하였으며, 모델의 하이퍼파라미터는 모두 20 에폭(epoch), 32 배치 사이즈, 1e-5 학습률, 0.1 드랍아웃(dropout), 768개 히든 사이즈(hidden size), 12개 히든 레이어(hidden layer)로 설정하였다.

3. Research Result

모델의 성능 비교에 있어 실제 발화자의 의도와 예측한 의도를 테스트 데이터에 대해 accuracy, f1, f1 weighted 총 세 가지 지표로 평가하였다.

모델별 한국어 Vocab 사이즈는 BERT의 경우 3,271개, KoBERT의 경우 8,002개, KoELECTRA의 경우 35,000개이다. 표1의 결과를 보면 Vocab 사이즈가 가장 큰 KoELECTRA 모델이 19개의 클래스 데이터에 대해 모든 지표에서 성능이 가장 우수한 것을 볼 수 있다. 한편, 발화자의 의도인 문의와 요청의 가장 큰 특징은 ‘물음표’를 제거하고 학습한 결과 그렇지 않은 경우보다 Accuracy는 0.04, f1은 0.09이 감소하였다.

Table 1. 실험 결과

클래스 수 및 구분	모델	Accuracy	f1	f1 weighted
19개	Bert	0.68	0.52	0.65
19개	KoBert	0.81	0.70	0.79
19개 (물음표 제거)	Ko Electra	0.77	0.61	0.75
19개	Ko Electra	0.81	0.70	0.80

IV. Conclusions

자연어 처리 분야의 BERT 기반 사전학습 모델은 일반적으로 학습 코퍼스의 양과 Vocab 사이즈가 클수록 성능이 향상된다. 본 연구에서는 이를 발화자 의도 분류 문제에 적용하여 사전학습 모델 성능을 비교함으로써 Vocab 사이즈가 큰 사전학습 모델이 우수한 성능을 보이는 것을 확인했다. 또한 발화자 의도 중 문의와 요청은 ‘물음표’ 문장 부호가 중요한 역할을 한다는 것도 확인했다. 향후 연구로는 발화의 의도를 보다 명확하게 파악하기 위해 한 문장 단위가 아닌, 문맥 단위 데이터를 사전학습 모델로 학습해봄으로써 모델 성능을 향상시킬 것으로 기대한다.

REFERENCES

- [1] Ji-eun Kim, Jin-ye Jang, Shin-Sim, Hye-dong Jeong, and Min-young Jeong. "Comparison of Deep Learning-based Intent Classification Model Performance for Conversational System Building". Proceedings of the Korean Information Science Association, p738-740. 2020.07.
- [2] Devlin, J., Chang, M.-W., Lee, K., and Toutanova, K. Bert: Pretraining of deep bidirectional transformers for language understanding. arXiv preprint arXiv:1810.04805, 2018.
- [3] Clark, K., Luong, M.-T., Le, Q. V., and Manning, C. D. ELECTRA: Pre-training text encoders as discriminators rather than generators. In International Conference on Learning Representations, 2020.
- [4] A. Bhargava, A. Celikyilmaz, D. Hakkani-Tur, R. Sarikaya, "EASY CONTEXTUAL INTENT PREDICTION AND SLOT DETECTION", 2013 IEEE International Conference on Acoustics, Speech and Signal Processing, pp. 8337-8341, 2013.