The 6th International Conference on Interdisciplinary research on Computer science, Psychology, and Education (ICICPE' 2022)
December 26-28, 2022. Pattaya, Thailand.

Joint learning method using intention classification and spacing to improve Korean SLU

Changwoo Chun[1] , Heuiseok Lim[1],

[1] Korea University, South Korea
{chun8629, limhseok}@korea.ac.kr

Abstract. A common method used for intent classification of short sentences consists of a language model (PLM) pre-trained as a text encoder and a classifier layer. Most tokenizers of PLM treat text input as split subwords while preserving whitespace. In this study, we observe that some information about spacing is lost during this tokenization process and propose a joint learning method using intent classification and spacing tasks as self-supervised learning. Space in input sentences is useful feature in the training process because there is no need of label. The proposed method contributes to the improvement of text classification performance, especially in Korean-based intention classification tasks.

Keywords: Intent classification, Joint learning, Spoken language understanding

1 Introduction

Intent classification is the field of natural language processing for classifying the intent of a speaker in a speech or dialogue sentences. [1]

Prior to the widespread use of speech recognition, the task of classifying the intent of an utterance did not receive much attention. Most natural language understanding(NLU) research focuses on understanding long sentences well, as long text data is very rich when collected from books, news articles, wiki, etc.

Recently, as voice recognition has been activated in AI speakers and cars, voice data from many users has begun to be collected. However, existing NLU models that focus on handling long sentences well often perform poorly for very short sentences.

Understanding long sentences is a very complex problem, while understanding short sentences well is a very ambiguous problem. Unlike human-to-human dialogue, machine-directed speech is unkind and frequently omits important words. [7]

When the functions provided by voice recognition were few, the problem was not serious. This is because even spitting out keywords was treated as a basic domain bias. However, as the range of functions that can be provided by voice recognition is gradually expanding, and people are freely speaking out of the set command set, domain bias is no longer a good solution. [4][5]

In this study, we propose an effective method for classifying the intent of a short sentence. This is a method to learn intention classification and spacing at the same time. Space information in Korean is very useful, but the tokenizer commonly used in

Pre-trained Language Model(PLM) does not convey space information sufficiently. And because the number of tokens used in short sentences is small, features are lacking to represent meaning of a sentence. The proposed method allows the model to distinguish important word boundaries in sentences by simultaneously learning space information, and improves intention classification performance by using additional features. In addition, space information does not require a separate labeling operation, enabling self-supervised learning.

.

2 Proposed Method

The proposed method uses space information. Tokens as a morphemed-subword [10] are input of model and an intent of sentence and space information are target for prediction. Space information is expressed as a list with values of 0 and 1. If a token is split by a morpheme analyzer or tokenizer, it has a value of 0, and if a token is a word separated by space in the first place, it has a value of 1.

For example, the sentence "Turn on the air conditioner" is tokenized to ["turn", "on", "the", "air", "cond", "##tion", "##er"] In this case, the space information is expressed as [1,1,1,1,0,0,0]. Transformer-based language model uses WPM or SPM-based tokenizer as above, and each token is input with one-hot encoding. Therefore, character information such as "##", which a person displays later to restore the original sentence, cannot be known from the model point of view. Therefore, inputting space information together can express the original meaning of the sentence more accurately.

In the mobility domain, words expressing destination information or addresses for movement are frequently used. In the realm of proper nouns, space information is ignored differently from grammar or used as a means of expressing specific meaning. For example, in hashtags or business names, the method of attaching all words without spaces is widely used. In this case, information that there are no spaces at all becomes very useful when analyzing the meaning of a sentence.

3 Experiments

All experiments were conducted based on the Korean model of BERT [3][10]. A method of not using space information, a method of using space information as input values at the same time, and a method of simultaneously predicting space information were compared through experiments.
Since it was an experiment on a test set with high performance in the first place, there was no significant performance difference between the comparison groups.

Through other studies, it can be seen that when learning the intent, the performance is improved by simultaneously learning the slot. [2][6][8][9] However, for this, data that is tagged with an intent as well as a slot is required.

Although the proposed method does not reach the level of learning slots at the same time, it can be confirmed that the performance is slightly improved compared to the method of learning only intents. Because the performance difference is small, it is necessary to perform additional experiments with statistical significance tests or other datasets.

Method	Algorithm	Accuracy	F1
Intent (baseline)	BERT-base	0.979	0.968
Intent (with space feature)	BERT-base	0.979	0.969
Intent-space joint learning	**BERT-base**	**0.980**	**0.972**
Intent-slot joint learning	BERT-base	0.992	0.990

Acknowledgments. This research was supported by the MSIT(Ministry of Science and ICT), Korea, under the ITRC(Information Technology Research Center) support program(IITP-2022-2018-0-01405) supervised by the IITP(Institute for Information & Communications Technology Planning & Evaluation). This research was supported by Basic Science Research Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Education(NRF-2021R1A6A1A03045425). Heuiseok Lim† is a corresponding author.

References

1. Won Ik Cho et al., "Speech Intention Understanding in a Head-final Language: A Disambiguation Utilizing Intonation-dependency", arXiv preprint arXiv:1811.04231, 2019.
2. Seunggyu Han, Heuiseok Lim, "Development of Korean dataset for joint intent classification and slot filling", Journal of the Korea Convergence Society Vol. 12. No. 1, pp. 57-63, 2021
3. Jacob Devlin et al., "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding", arXiv preprint arXiv:1810.04805, 2018
4. Y. Kim., "Convolutional neural networks for sentence classification", arXiv preprint, arXiv:1408.5882.(2014)
5. K. Yao et al., "Spoken language understanding using long short-term memory neural networks", 2014 IEEE Spoken Language Technology Workshop (SLT), 189–194.(2014)
6. P. Xu and R. Sarikaya, "Convolutional neural network based triangular crf for joint intent detection and slot filling", Automatic Speech Recognition and Understanding (ASRU), 2013 IEEE Workshop on. IEEE, 2013, pp. 78–83.
7. G. Tur, D. Hakkani-Tur, L. Heck, and S. Parthasarathy, "Sentence simplification for spoken language understanding", in Acoustics, Speech and Signal Processing (ICASSP), 2011 IEEE International Conference on. IEEE, 2011, pp. 5628–5631.
8. Bing Liu, Ian Lane, "Attention-Based Recurrent Neural Network Models for Joint Intent Detection and Slot Filling", arXiv:1609.01454v1 [cs.CL] 6 Sep 2016
9. Q. Chen, Z. Zhuo, & W. Wang., "Bert for joint intent classification and slot filling". arXiv preprint, arXiv:1902.10909 (2019)
10. yeontaek, BERT with MECAB tokenizer for Korean text, https://github.com/yeontaek/BERT-MECAB-Korean-Model