https://doi.org/10.5392/JKCA.2019.19.11.278

Bi-LSTM-CRF 앙상블 모델을 이용한 한국어 공간 정보 추출

Korean Spatial Information Extraction using Bi-LSTM-CRF Ensemble Model

민태홍[*], 신형진[*], 이재성[**]
충북대학교 컴퓨터과학과[*], 충북대학교 소프트웨어학과[**]

Tae Hong Min(th.min@kt.com)[*], Hyeong Jin Shin(hjshin@cbnu.ac.kr)[*], Jae Sung Lee(jasonlee@cbnu.ac.kr)[**]

요약

공간 정보 추출은 자연어 텍스트에 있는 정적 및 동적인 공간 정보를 공간 개체와 그들 사이의 관계로 명확히 표시하여 추출하는 것을 말한다. 이 논문은 2단계 양방향 LSTM-CRF 앙상블 모델을 사용하여 한국어 공간 정보를 추출할 수 있는 심층 학습 방법을 제안한다. 또한 공간 개체 추출과 공간 관계 속성 추출을 통합한 모델을 소개한다. 한국어 공간정보 말뭉치(Korean SpaceBank)를 사용하여 실험한 결과 제안한 심층학습 방법이 기존의 CRF 모델보다 우수함을 보였으며, 특히 제안한 앙상블 모델이 단일 모델보다 더 우수한 성능을 보였다.

■ 중심어 : | 공간 정보 | 정보 추출 | 심층 학습 | LSTM-CRF |

Abstract

Spatial information extraction is to retrieve static and dynamic aspects in natural language text by explicitly marking spatial elements and their relational words. This paper proposes a deep learning approach for spatial information extraction for Korean language using a two-step bidirectional LSTM-CRF ensemble model. The integrated model of spatial element extraction and spatial relation attribute extraction is proposed too. An experiment with the Korean SpaceBank demonstrates the better efficiency of the proposed deep learning model than that of the previous CRF model, also showing that the proposed ensemble model performed better than the single model.

■ keyword : | Spatial Information | Information Extraction | Deep Learning | LSTM-CRF |

I. INTRODUCTION

Tracking spatial information such as location and motion in natural language is important for many applications including robotics, question answering systems and navigation systems. Capturing spatial information with annotation schemes has been attempted in several ways: SpatialML[1], Spatial Role Labeling: SpRL[2], and ISO-Space[3][4]. For example, a relation, on (book, table), can be annotated in the following sentence, where *se* is a spatial entity with trajector role, *ss* is a spatial signal representing a relation, *pl* is a place with landmark role.

* 이 논문은 2019년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구임 (No.2013-2-00131, 휴먼 지식증강 서비스를 위한 지능진화형 Wise QA 플랫폼 기술 개발)
접수일자 : 2019년 09월 20일
수정일자 : 2019년 10월 14일
심사완료일 : 2019년 10월 24일
교신저자 : 이재성, e-mail : jasonlee@cbnu.ac.kr

There is a [book]$_{se}$ [on]$_{ss}$ the [table]$_{pl}$

Spatial information extraction methods have also been developed alongside these annotation schemes, especially through shared tasks in SemEval using SpRL 2010, 2012, 2013 and ISO-Space 2015[2][5-7]. While most of recent methods attempted spatial information extraction based on SVM[8] and CRF[9][10], many deep learning approaches have been developed for similar tasks such as information extraction[11-13] and semantic role labeling[14].

Because there is a difference between spatial information extraction and other similar tasks described above, the deep learning methods developed for those similar tasks should be modified or re-invented for spatial information extraction. For example, relation classification, a subtask of information extraction, determines relation types between two given nouns among a set of relations[15] while spatial information extraction decides explicitly the relation word, such as 'on' in the previous example. Semantic role labeling (SRL) finds both a predicate and its arguments in a sentence. The predicates (relations) are mainly verbs in SRL[16] while they can be either prepositions or verbs in spatial relation extraction.

Mazalov et al. utilized a deep learning program developed for SRL to process SpRL annotated text, where spatial signals such as 'on' are used instead of predicates[17]. It was developed for English text and was based on CNN model that is widely used for image processing.

In this paper, we propose an RNN-based deep learning approach for spatial information extraction using a two-step bidirectional-LSTM-CRF ensemble model. We apply the ensemble model to spatial element extraction, spatial relation extraction, and their integration system with a pipeline approach. The models are developed for Korean text considering Korean specific features such as word-phrases and morphemes[10][18].

Section II describes the proposed deep learning model. Section III describes the element extraction and relation extraction models based on the proposed ensemble model. The experiment result and discussion follow in Section IV, and conclusion in Section V.

II. Two-Step Bidirectional LSTM CRF Ensemble Model

1. Bi-LSTM model

The CRF model has been proven effective for spatial information extraction in both English and Korean texts[4][10]. In this paper, we use a bidirectional-LSTM-CRF model with some modification to automatically extract features, of which basic structure was used for named entity extraction by[19].

LSTM is an RNN cell that has the advantage of learning long sequences. LSTM uses three gates: input gate (it), output gate (ot) and forget gate (ft)[20][21]. [Fig. 1] shows a LSTM cell.

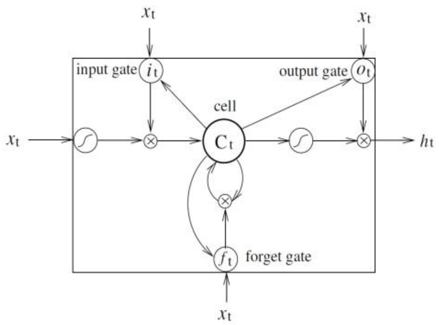

Fig. 1. A Long Short-Term Memory Cell[20]

We use the following LSTM implementation.

$$f_t = \sigma(W_{xf}x_t + W_{hf}h_{t-1} + W_{cf}c_{t-1} + b_f) \quad (1)$$
$$i_t = \sigma(W_{x_i}x_t + W_{hi}h_{t-1} + W_{ci}c_{t-1} + b_i) \quad (2)$$
$$g_t = \tanh(W_{xc}x_t + W_{hc}h_{t-1} + b_g) \quad (3)$$
$$c_t = f_t \odot c_{t-1} + i_t \odot g_t \quad (4)$$
$$o_t = \sigma(W_{xo}x_t + W_{ho}h_{t-1} + W_{co}c_t + b_o) \quad (5)$$
$$h_t = o_t \odot \tanh(c_t) \quad (6)$$

Equation 1 determines how much of the previous value to be used. Equations 2 and 3 determine how much of the current value to be used. Equation 4, which combines equations 1, 2 and 3, is the output of an LSTM cell, where ⊙ is a Hadamard product meaning the element-wise multiplication of two matrices. Equations 5 and 6 transmit information of the current LSTM cell to the next LSTM cell.

We also use Bi-LSTM, which takes bi-directional input, using two LSTMs for forward and backward LSTMs. [Fig. 2] is a model of Bi-LSTM[20].

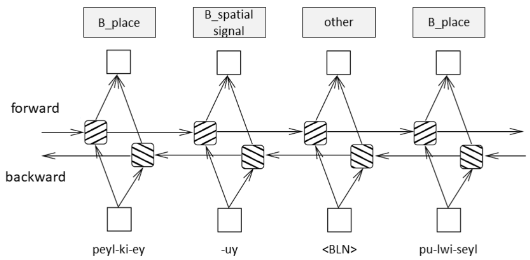

FIg. 2. A Bi-LSTM model (〈BLN〉 means a blank)

2. Bi-LSTM-CRF model

Finally, we use stacked LSTM with two LSTM layers. The input of the second LSTM uses the output of the first LSTM. The output of stacked Bi-LSTM is concatenated and used as the features of CRF. [Fig. 3] is the spatial elements extraction model with stacked Bi-LSTM.

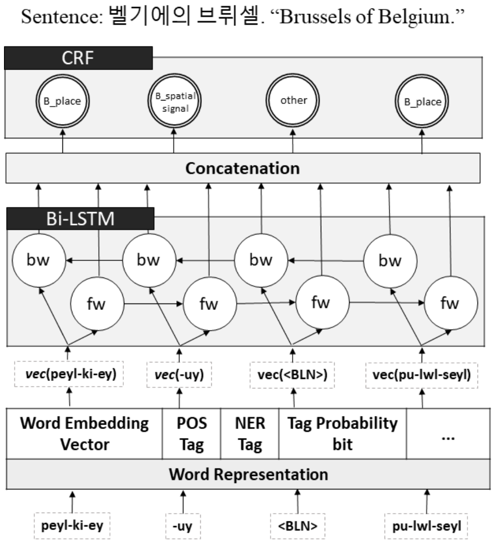

Fig. 3. Bi-LSTM-CRF model for a sample Korean phrase: 벨기에의 브뤼셀 "peyl-ki-ey-uy pu-lwi-seyl"

Korean morphemes are input as vectors in word representation layer in the model. The word representation layer is connected with bidirectional LSTM layer and the output of both directions is concatenated into new vectors. The concatenated vectors are used as input for CRF, producing the final result with IOB prefixed spatial tags.

3. Bi-LSTM-CRF ensemble model

The Bi-LSTM-CRF ensemble model was further developed into a two-step ensemble model, alleviating the data sparseness problem primarily resulting from small training data sets. [Fig. 4] shows the model: input is processed by 5 Bi-LSTM-CRF models in the first step, and their outputs are summed into the input for another Bi-LSTM-CRF model in the second step. Randomness in the initial weight and drop out settings causes each Bi-LSTM-CRF model to be trained with different weights, creating five different models. Preliminary test has shown that the ensemble model is more effective than the simple model so that the two-step ensemble model is applied to both element and relation extractions.

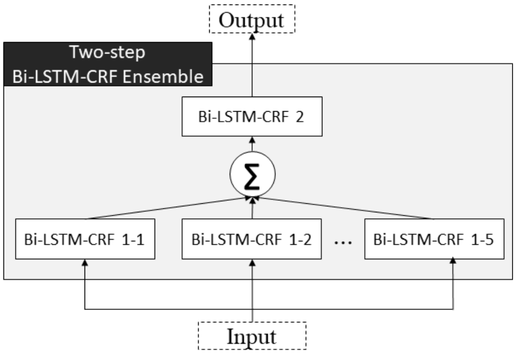

Fig. 4. Two-step Bi-LSTM-CRF ensemble model

III. Spatial Information Extraction

1. Overview

Spatial information is extracted in two phases; elements extraction and relation extraction. [Fig. 5] shows the overview of the proposed system. The relations can be extracted either from the previous phase's results or from the sentences of golden annotation. The details are described in the following subsections.

2. Spatial Elements Extraction

Spatial elements extraction targets the seven elements defined in ISO-Space: place, path, spatial entity, spatial signal, motion, motion signal, and measure. Elements are tagged with IOB tags (total 15 labels) for sequential labeling with the Bi-LSTM-CRF model: two BI tags for each of the seven elements and one O tag for unrelated words.

As word representation of input is important for model performance, it was carefully determined to use the following features vectors with four types of information, forming 196 dimensions:

- Word embedding vector: 100 dimensions of Korean morpheme embedding, produced using fastText[22], pre-trained with Korean SpaceBank corpus of 17K word-phrases and Sejong raw corpus of 10M word-phrases[23][24]
- Morpheme-POS tag: 46 one-hot vector dimensions, including 45 Part-of-Speech tags for Korean morphemes plus one sentence boundary marking tag[25].
- Named entity tag: 35 one-hot vector dimensions representing the highest and the second highest classifications of location, artifacts, and quantity defined in[26].
- Tag probability bit: 15-bit vectors for each morpheme tag, appearing more than k times in the training data. (In this experiment, k is 3)

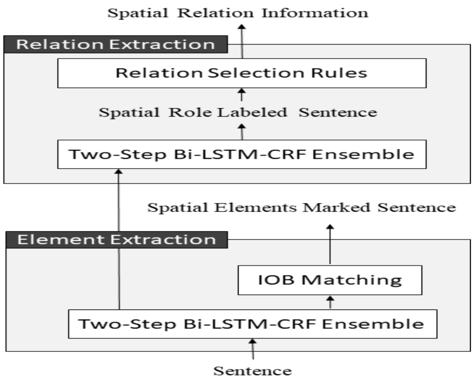

Fig. 5. Overview of proposed spatial information extraction system (integrated model)

3. Spatial Relation Extraction

Spatial relation extraction connects previously retrieved spatial elements with four relations: qualitative spatial link (qsLink), orientation link (oLink), movement link (moveLink), and measure link (mLink). qsLink, oLink, and mLink are represented in the triple format 〈trajector, landmark, trigger〉, where trajector represents the subject of the relation, landmark represents the base or ground of the relation, and trigger explains the relation itself. moveLink is represented in 7-tuples in SpaceEval; however, we extracted only the triple 〈mover, goal, trigger〉 which has typically been adopted in the previous research[9][10].

Relations were extracted in two steps: first, spatial roles (trajectors-landmarks or movers-goals) are determined by two-step Bi-LSTM-CRF ensemble model; second, trajectors and landmarks (movers and goals) are connected with triggers to create triple relations by the relation selection rules. Each relation is extracted independently so that an element could be involved in multiple relations. [Fig. 6] shows an example, where the element "syo-phing-mol" involves in both qsLink and moveLink relation. When an element plays both the trajector and landmark roles in single relation, it is marked as "traLand" in role determination process.

Word representation for relation extraction is composed of seven information types, forming 295 dimensions. This includes the word embedding, morpheme-POS tag, and tag probability bit vectors used in element extraction. The following vectors are also included to handle contextual dependency

- Spatial element tag: 15 dimensions of one-hot vector representing spatial element tags extracted in the first step of spatial information extraction.
- Dependency label: 19 dimensions of one-hot vector representing 8 grammatical tags, 7 functional tags, and etc., described in[27].
- Head's dependency label: 19 dimensions of one-hot vector representing the dependency label of the head of the dependency tree.
- Main morpheme-POS tag of head: 46 dimensions of one-hot vector representing the main morpheme-POS tag of the head of the dependency tree.

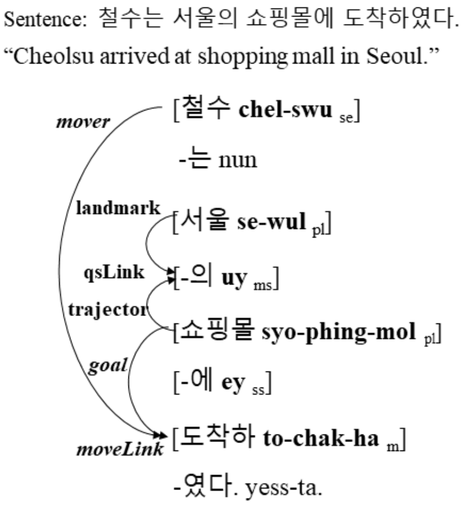

Fig. 6. Example of one element involving multiple relations

4. Relation Selection Rules

In this experiment, we set the selection window size to 17 for qsLink and oLink, 9 for moveLink, and 10 for mLink respectively. All the possible combinations of relations are tried in order to find the legitimate ones within the respective window sizes and with the following rules.

- Only spatial signal, motion and measure can be a trigger. And only place, path, and spatial entity can be a trajector and landmark (or mover and goal).

- A traLand is related to the closest trigger in the dependency tree in mLink, but a traLand is related to all triggers within the window size in qsLink and oLink.
- Null arguments are allowed in moveLink when sentences do not contain either mover or goal.

IV. Experiment

1. Environment

We used Korean SpaceBank version 2.1 (minor error correction of version 2.0), and [Table 1] shows the statistics of the corpus[23]. This corpus is used for our 5-fold-cross validation test, which divides the data at 4:1 ratio for training and testing.

Table 1. Number of tags in the testing corpus

Elements	# tags	Relations	# tags
place	4,922	qsLink	1,159
path	280	oLink	565
s.Entity	396	moveLink	305
motion	254	mLink	345
m.Signal	249	total	2,374
s.Signal	1,245		
measure	282		
total	7,628		

System configuration for the experiment is shown in [Table 2]. And all the models were tested with the following hyper-parameters as shown in [Table 3].

Table 2. System configuration

CPU	intel i7-7700
GPU	Geforce 1080Ti 2ea
RAM	48GB
O/S	UBUNTU 16.04
Development language	python 2.7

Table 3. Model hyper-parameters

LSTM cell size	256 dimension, 2 layers
loss function	cross-entropy
optimizer	Adam optimizer
dropout rate	0.2
learning rate	0.01
batch size	64

2. Result

The model of Kim and Lee[10] is used as a baseline model, which shows the state-of-art performance for Korean spatial information extraction. It is based on CRF model with manually provided features and it is retested with new Korean test data for comparison.

We performed mainly three evaluation tasks: spatial element extraction using raw text data (for short, task *ele*), spatial relation extraction using given elements (for short, task *rel*), and spatial relation extraction using raw text data, integrating both element extraction task and relation extraction task (for short, task *irel*, which is the ultimate goal of spatial information extraction). These correspond to the task definitions 1.b, 2.b, and 1.d respectively in Pustejovsky et al. (2015). One substitutional evaluation task is added for the model of Kim and Lee[10], which evaluated spatial relation extraction from given elements with attributes (for short, task *arel*), which corresponds to 3.a of the SemEval task definition. Kim and Lee[10] uses the attributes of the spatial signal to distinguish whether the spatial signal is a trigger of qsLink or a trigger of oLink.

The test result of spatial element extraction (task *ele*) is shown in [Table 4]. We used relaxed evaluation, which accepts the partial extent matches as correct. These results demonstrate that the proposed model has improved performance than the baseline in averages, indicating the deep learning model worked better than the CRF baseline model. In particular, the performance of path and spatial entity extraction is better than the baseline model. These spatial elements are one of the causes of performance degradation in the baseline model, which are mainly nouns and require contextual information to distinguish them from places: place element is dominant as shown in [Table 1] and causes data skewness problem.

Table 4. Spatial element extraction: task *ele* (F1, %)

Elements	Base	Prop	diff
place	94.0	92.7	−1.3
path	54.0	72.8	18.8
s.Entity	38.5	42.6	4.1
motion	51.9	66.6	14.7
m.Signal	62.5	60.4	−2.1
s.Signal	82.2	82.3	0.1
measure	88.5	92.0	3.5
mic avg	84.2	86.5	2.3

Relation extraction task of the proposed model is processed in two stages: role determination and relation selection as shown in [Fig. 3]. [Table 5] shows the results of spatial relation role extraction. [Table 6] compares the proposed model with the baseline model for spatial information extraction, showing that all relations demonstrated improved performance than the baseline model. (We omitted the mLink result only because the baseline model did not report the performance result in the paper.) Especially, the precision of proposed model increased by 29.8% point in micro average compared with baseline model, while the recall decreased by 6.9% point, which leads to performance improvements in F1. Note that the task *arel* is easier than the task *rel* because the link type of spatial signal is determined or given before the experiment in the task *arel*. This fact also strongly supports that the proposed model is much better than the baseline model.

Table 5. Spatial relation role determination (%)

Rel	Role	F1	Rel	Role	F1
qs Link	trajector	78.1	move Link	mover	85.3
	landmark	84.2		goal	65.7
	trigger	94.7		trigger	91.8
	traLand	57.3		-	-
	mic avg	84.7		mic avg	81.9
o Link	trajector	64.7	m Link	trajector	63.6
	landmark	71.4		landmark	64.3
	trigger	85.2		trigger	92.7
	traLand	26.5		traLand	63.7
	mic avg	72.8		mic avg	75.6

Table 6. Spatial relation extraction performance of baseline and proposed model: task *arel* and *rel* (%)

Relation	Base(*arel*)			Prop(*rel*)		
	Pre	Re	F1	Pre	Re	F1
qsLink	47.1	55.7	51.1	75.3	51.1	60.9
oLink	25.4	50.4	35.9	67.7	44.3	53.6
moveLink	22.4	62.2	33.0	40.0	42.4	41.2
mic avg	35.9	55.0	43.5	65.7	48.1	55.6

The ensemble model performed better than the single model both for spatial element extraction task and for spatial relation extraction task, as shown in [Table 7] and [Table 8].

Table 7. Comparison of the ensemble model with the single model for spatial element extraction task (F1, %)

Elements	single	ensemble	diff
place	91.8	92.7	0.9
path	70.8	72.8	2.0
s.Entity	36.7	42.6	5.9
motion	59.7	66.6	6.9
m.Signal	58.7	60.4	1.7
s.Signal	79.8	82.3	2.5
measure	89.6	92.0	2.4
mic avg	84.7	86.5	1.8

Table 8. Comparison of the ensemble model with the single model for spatial relation extraction (F1, %)

Relation	single	ensemble	diff
qsLink	39.3	42.8	3.5
oLink	43.2	47.4	4.2
moveLink	27.0	25.2	-1.8
mLink	40.7	42.3	1.6
micro avg	**38.6**	**41.2**	**2.6**

[Table 9] shows the result of integrated test (task *irel*), compared with the method using annotated spatial element (task *rel*). The integrated task uses unannotated test data, and goes through both spatial element extraction and relation extraction phases. It is natural that the performance of the integrated system decreases because of error propagation from previous phase. We do not compare the *irel* result with the baseline model because the baseline model is not implemented for the *irel* task. However, we can conjecture that the proposed model's *irel* task performance will be better than the baseline model's because the proposed model performed better in both element extraction and relation extraction as we have seen before.

Table 9. Spatial relation extraction using annotated and un-annotated corpus: task *rel* and *irel* (F1, %)

Relation	task *rel*	task *irel*	diff
qsLink	60.9	42.8	-18.1
oLink	53.6	47.4	-6.2
moveLink	41.2	25.2	-16.0
mLink	49.9	42.3	-7.6
micro avg	**54.8**	**41.2**	**-13.6**

All the evaluations in this paper are done with five-fold cross validation test.

V. Conclusion

We have proposed a two-step Bi-LSTM-CRF ensemble model for Korean spatial information extraction. The proposed model is the first application of RNN based deep learning model to spatial information extraction, as far as we know. The model adopts ensemble technique, and proposed word representations suitable for Korean spatial information, and a sequence labeling method for multiple relation extraction.

As we developed our models for Korean version, we mainly compared with previous state-of-art Korean models using the manually featured CRF model. We compared the performance of two factors for comparison: spatial element and spatial relation alone. An experiment using the Korean SpaceBank showed that our proposed models performed better than the baseline models. We also implemented and tested the integrated relation extraction model for end to end relation extraction.

We used Korean word embeddings trained with fastText for our Bi-LSTM-CRF model in this paper. As a further research, we will use for our model the contextual word vectors such as ELMo[27] and BERT[28], which show much better performance in deep learning applications.

참 고 문 헌

[1] I. Mani, J. Hitzeman, J. Richer, D. Harris, R. Quimby, and B. Wellner, "SpatialML: Annotation Scheme, Corpora, and Tools," In LREC, 2008.

[2] P. Kordjamshidi, M. F. Moens, and M. van Ctterlo, "Spatial role labeling: Task definition and annotation scheme," In Proceedings of the Seventh conference on International Language Resources and Evaluation (LREC'10), European Language Resources Association (ELRA), pp.413-420, 2010.

[3] ISO-24617-7:2014, Language resource management - part 7: Spatial information (ISOspace).

[4] J. Pustejovsky, P. Kordjamshidi, M. F. Moens, A. Levine, S. Dworman, and Z. Yocum, "SemEval-2015 Task 8: SpaceEval," In Proceedings of the 9th International Workshop on Semantic Evaluation (SemEval 2015), pp.884-894, 2015.

[5] P. Kordjamshidi, S. Bethard, and M. F. Moens, "SemEval-2012 task 3: Spatial role labeling," Proceedings of the First Joint Conference on Lexical and Computational Semantics-Volume 1: Proceedings of the main conference and the shared task, and Volume 2: Proceedings of the Sixth International Workshop on Semantic Evaluation, Association for Computational Linguistics, 2012.

[6] O. Kolomiyets, P. Kordjamshidi, M. F. Moens, and S. Bethard, "Semeval-2013 task 3: Spatial role labeling," Second Joint Conference on Lexical and Computational Semantics (* SEM), Volume 2: Proceedings of the Seventh International Workshop on Semantic Evaluation (SemEval 2013), Vol.2, 2013.

[7] J. Pustejovsky, P. Kordjamshidi, M. F. Moens, A. Levine, S. Dworman, and Z. Yocum, "SemEval-2015 Task 8: SpaceEval," In Proceedings of the 9th International Workshop on Semantic Evaluation (SemEval 2015), pp.884-894, 2015.

[8] H. Salaberri, O. Arregi, and B. Zapirain, "IXAGroupEHUSpaceEval:(X-Space) A WordNet-based approach towards the automatic recognition of spatial information following the ISO-Space annotation scheme," In Proceedings of the 9th International Workshop on Semantic Evaluation (SemEval 2015), pp.856-861, 2015.

[9] E. Nichols and F. Botros, "SpRL-CWW: Spatial relation classification with independent multi-class models," In Proceedings of the 9th International Workshop on Semantic Evaluation (SemEval 2015), 2015.

[10] B. Kim and J. S. Lee, "Extracting Spatial Entities and Relations in Korean Text," In Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: Technical Papers, pp.2389-2396, 2016.

[11] P. N. Golshan, H. R. Dashti, S. Azizi, and L. Safari, "A Study of Recent Contributions on Information Extraction," arXiv:1803.05667, 2018.

[12] M. Miwa and Y. Sasaki, "Modeling joint entity and relation extraction with table representation," In Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP), pp.1858-1869, 2014.

[13] M. Miwa and M. Bansal, "End-to-end relation extraction using lstms on sequences and tree structures," arXiv preprint arXiv:1601.00770, 2016.

[14] L. He, K. Lee, M. Lewis, and L. Zettlemoyer, "Deep semantic role labeling: What works and what's next," In Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), Vol.1, pp.473-483, 2017.

[15] I. Hendrickx, S. N. Kim, Z. Kozareva, P. Nakov, D. Ó Séaghdha, S. Padó, and S. Szpakowicz, "Semeval-2010 task 8: Multi-way classification of semantic relations between pairs of nominals," In Proceedings of the Workshop on Semantic Evaluations: Recent Achievements and Future Directions, Association for Computational Linguistics, pp.99-99, 2009.

[16] L. Marquez, X. Carreras, K. C. Litkowski, and S. Stevenson, "Semantic role labeling: an introduction to the special issue," Computational Linguistics, Vol.34, No.2, pp.145-159, 2008.

[17] A. Mazalov, B. Martins, and D. Matos, "Spatial role labeling with convolutional neural networks," Proceedings of the 9th Workshop on

Geographic Information Retrieval, ACM, 2015.
[18] B. Kim, M. Y. Kang, and J. S. Lee. "Issues in spatial information annotation in Korean texts," In Proceedings of 2016 International Conference on Big Data and Smart Computing (BigComp), 2016.
[19] G. Lample, M. Ballesteros, S. Subramanian, K. Kawakami, and C. Dyer, "Neural architectures for named entity recognition," arXiv preprint arXiv:1603.01360, 2016.
[20] S. Hochreiter and J. Schmidhuber, "Long short-term memory," Neural computation, Vol.9, No.8, pp.1735-1780, 1997.
[21] Z. Huang, W. XU, and K. Yu, "Bidirectional LSTM-CRF models for sequence tagging," arXiv preprint arXiv:1508.01991, 2015.
[22] P. Bojanowski, E. Grave, A. Joulin, and T. Mikolov, "Enriching word vectors with subword information," arXiv preprint arXiv:1607.04606, 2016.
[23] CBNU(Chungbuk National University), Language and Knowledge Engineering Lab, *Korean SpaceBank v2.0 Guideline*, 2017.
[24] NIKL(National Institute of Korean Language), *21st century Sejong project final result*, revised edition. 2011.
[25] TTAK.KO-11.0010/R1, "Part-of-Speech Tag Set for Morphological Annotation of Written Texts," 2015.
[26] TTAK.KO-10.0852, "Tag Set and Tagged Corpus for Named Entity Recognition," 2015.
[27] TTAK.KO-10.0853, "Dependency Tag Sets and Dependency Relation Establishment Methods for Constructing Dependency Tagged Corpora," 2015.
[28] M. E. Peters, M. Neumann, M. lyyer, M. Gardner, C. Clark, K. Lee, and L. Zettlemoyer, "Deep contextualized word representations," arXiv preprint arXiv:1802.05365, 2018.
[29] J. Devlin, M. W. Chang, K. Lee, and K. Toutanova, "Bert: Pre-training of deep bidirectional transformers for language understanding," arXiv preprint arXiv:1810.04805, 2018.

저 자 소 개

민 태 홍(Tae Hong Min) 정회원

- 2012년 ~ 2017년 : 충북대학교 컴퓨터교육과(학사)
- 2017년 ~ 2019년 : 충북대학교 컴퓨터과학과(석사)
- 2019년 ~ 현재 : KT융합기술원, 연구원

〈관심분야〉 : 자연어처리, 기계학습

신 형 진(Hyeong Jin Shin) 준회원

- 2014년 ~ 2018년 : 충북대학교 컴퓨터공학과(학사)
- 2018년 ~ 현재 : 충북대학교 컴퓨터과학과 석사과정

〈관심분야〉 : 자연어처리, 정보추출, 딥러닝

이 재 성(Jae Sung Lee) 정회원

- 1979년 ~ 1983년 : 서울대학교 컴퓨터공학과(학사)
- 1983년 ~ 1985년 : KAIST 전산학과(석사)
- 1985년 ~ 1988년 : 큐닉스 과장
- 1988년 ~ 1993년 : 마이크로소프트 차장
- 1995년 ~ 1999년 : KAIST 전산학과 (박사)
- 1999년 ~ 2000년 : ETRI 지식정보검색연구팀 팀장
- 2000년 ~ 현재 : 충북대학교 컴퓨터 교육과, 소프트웨어학과 교수

〈관심분야〉 : 자연어처리, 정보검색, 기계학습