관용표현 중의성 해소를 위한 다층위 임베딩 연구

박서윤○, 강예지, 강혜린, 장연지◇, 김한샘●

연세대학교 언어정보학협동과정, 국립국어원◇, 언어정보연구원●
{seoyoon.park○, yjkang5009, hyerink, yeonji3547◇, khss●}@yonsei.ac.kr

Embedding with different levels for idiom disambiguation

Seo-Yoon Park○, Ye-Jee Kang, Hye-Rin Kang, Yeon-Ji Jang◇, Han-Saem Kim●
Interdisciplinary Graduate Program of Linguistics and Informatics, Yonsei University
National Institute of Korean Language◇
Institute of Language and Information Studies, Yonsei University●

요 약

관용표현 중에는 중의성을 가진 표현이 많다. 즉 하나의 표현이 맥락에 따라 일반적 의미와 관용적 의미 두 가지 이상으로 해석될 가능성이 있어 이런 유형의 관용표현을 중의성 해소 없이 자연어 처리 태스크에 적용할 경우 문제가 발생하게 된다. 본 연구에서는 관용표현의 특성인 중의성과 더불어 '관용표현은 이미 사용자의 머릿속에 하나의 토큰으로 저장되어 있다'라는 'Idiom Principle'을 바탕으로 관용표현에 대해 각각 표면형, 단순 단일 토큰형, stemming 단일 토큰형 층위의 임베딩을 만들어 관용표현 분류 연구를 진행하였으며, 실험 결과 표면형 및 stemming을 적용하지 않은 단순 단일 토큰으로 학습하는 것보다, stemming을 적용한 후 단일 토큰으로 학습하는 것이 관용표현의 중의성 해소에 유의미한 효과가 있음을 확인하였다.

주제어: 관용표현, 중의성 해소, idiom principle, Word2Vec, FastText

1. 서론

언어적 직관이 부족한 기계의 경우 사람처럼 관용표현의 의미를 정확하게 이해하기 힘들다. 관용표현의 대부분이 직설적 의미(literal)와 관용적 의미(idiomatic) 모두를 가진 중의적인 표현이기 때문이다.

가. 그는 조심스럽게 **문을 열어** 택배를 확인했다.
나. 저 방앗간이 **문을 연** 지도 벌써 10년 째다.

'가' 의 경우 사람이 드나들 수 있는 곳에 설치된 문을 열었다는 뜻으로 글자 그대로 직설적 의미로 쓰인 경우이다. 반면 '나' 의 경우 '문을 열다' 가 '영업을 하다' 란 뜻으로 쓰였으므로 관용적 의미로 쓰인 경우에 속한다. 이처럼 관용표현은 같은 표현이 두 개 이상의 의미를 가진 중의적 언어 단위로, 이를 실제 자연어 처리 태스크에 적용할 경우 중의성 해소 문제가 발생할 수밖에 없다.

이를 해결하기 위한 기초 연구로는 문맥 내 표현이 관용적 의미로 쓰였는지 혹은 문자적 의미로 쓰였는지 분류하는 태스크[1,2]나 혹은 문맥 안 표현이 관용표현인지 아닌지를 탐지(detection)하는 태스크[3]가 있다.

또한 관용표현은 인지하거나 실제로 표현을 사용할 때 '한 덩어리' 로 인식되는 특징이 있다. 이는 관용표현이 '습관적으로' 사용되기 때문이다[4]. 또한 [5]에서는 이러한 관용표현의 특징을 'Idiomatic Principle' 로 설명하였다. 'Idiom Principle' 이란 '미리 구성되어 있는(semipreconstructed) 표현들은 이미 언어 사용자의 머릿속에 존재한다' 라는 원칙이다. 즉 단어 이상의 표현들은 이미 언어 사용자들의 머릿속 사전(mental lexicon)에 하나의 토큰 내지 원형으로 저장되어 있다[6].

본고는 관용표현의 중의성 해소에 기여하고자 관용표현에 대한 분류 태스크 연구를 수행하였다. 특히 관용표현의 특성인 '습관적 사용' 과 더불어 '관용표현은 언어 사용자의 머릿속에 이미 한 토큰으로 인식된다' 라는 내용을 바탕으로 한 'idiom principle' 을 반영하여 임베딩 생성 시 관용표현을 '_' 로 이은 단일 토큰으로 처리하여 연구를 진행하였다.

이를 위해 2장에서는 연구의 배경이 되는 관련 연구를 소개하였으며, 3장에서는 연구에 사용한 데이터셋인 '관용표현 레이블링 말뭉치' 와 데이터셋 전처리를 다루었다. 4장은 관용표현 분류에 대한 실험과 더불어 결과를 서술하였고 5장은 전체 연구에 대한 결론을 내렸다.

2. 관련 연구

2.1. 관용표현 분류 연구

관용표현은 '표현에 드러난 단어들의 의미의 합으로 유추될 수 없는, 습관적으로 쓰이는 제 3의 의미를 가진 표현' 으로 정의할 수 있다.[4,7,8] 또한 관용표현은 표현 하나에 문자 그대로의 의미를 가지는 '직설적 의미'와 관용적으로 쓰이는 '관용적 의미' 두 가지를 가짐으로써 중의성이 발생한다. 이러한 특성 때문에 관용표현의 중의성을 고려하면서 자연어 처리 태스크를 수행하는 것은 매우 어려우며 아직까지도 해결되지 않는 자연어 처리 분야이다.

딥러닝 시대 이전 관용표현의 중의성을 고려한 전산 언어학적 연구로는 [9,10,11]을 들 수 있다. 우선 관용표현의 언어적 특성에 집중하여 진행된 연구로는 'type-based' 접근법을 들 수 있다[9,10]. 이 접근법에서는 어휘적 고정성(lexical fixedness) 나 통사적 고정성(syntactic fixedness)을 중심으로 관용표현의 중의성을 해소하고자 하였다. 그러나 'type-based' 연구의 경우 표현과 주변 문맥(context)을 고려하지 못한다는 한계가 존재하였고, 이를 고려하기 위해 'token-based' 연구가 이루어졌다[11]. 'Token-based' 의 대표적인 연구 방법론으로는 'Cohesion Graph' 가 있으며, 관용표현은 주변 문맥과의 응집성이 떨어지는 것을 전제로 한다. 후에 [12]에서는 'Cohesion Graph' 와 딥러닝 임베딩 기법을 결합하기도 하였다.

딥러닝 이전 시기 국내에서도 관용표현의 중의성 해소를 위한 연구들이 진행되었다[13,14]. [13]의 경우 관용표현의 자연어 처리를 고려하여 관용표현의 중의성 해소를 위한 알고리즘을 고안하였다. 알고리즘은 관용표현의 언어적 특성을 고려한 판별기제를 바탕으로 고안되었으며, 격틀, 수식어구 여부, 논항 자질, 동사 활용 형태, 조사 생략 여부 등을 반영하였다. 딥러닝 이전에는 통계기반 기계 번역(Statistical Machine Translation)을 사용하였는데, [14]에서는 다어절변환단위(MWTU; Multi - Word Translation Unit)을 바탕으로 숙어, 구어적 관용구의 의미를 타겟 언어와 맵핑한 MWTU를 사용해 한-일 번역의 품질을 향상하였다.

딥러닝이 소개된 이후에는 딥러닝 기법을 활용한 관용표현 중의성 해소 연구가 이루어졌다. [3]의 경우 Word2Vec과 tf-idf 기법을 활용하여 관용표현과 일반표현 간의 분류 태스크를 수행하였다. 또한 [1]에서는 형태론적으로 발달한 프랑스어, 바스크어에 대해 관용표현 분류 태스크를 수행하였다. 이 때 Word2Vec과 FastText가 활용되었는데, subword 생성을 특성으로 하는 FastText가 문맥 내 관용표현의 경계를 더 잘 포착하여 Word2Vec에 비해 강건한 결과를 나타냈다. 문맥 반영 임베딩을 만들어내는 ELMo를 활용하여 진행된 연구로는 [2]가 있다. [2]에서는 Word2Vec, FastText, ELMo를 영어, 독일어 관용표현 분류 태스크에 적용하였으며 문맥을 반영하는 ELMo의 성능이 가장 높았다.

2.2. Idiom Principle을 반영한 관용표현 분류 연구

원형을 전제하고 진행된 연구로는 [15]를 들 수 있다. [15]에서는 관용표현이 일반적 의미와 관용적 의미로 사용될 때의 형태가 다른 것을 포착하여 영어의 관용표현을 레마화(lemmatization)하였다. 이를 Supervised learning을 사용하여 분류 태스크를 수행한 결과, 표면형(surface form)을 사용했을 때보다 성능 향상이 뚜렷하였다. 또한 BERT를 분류 태스크에 사용하여 문맥 임베딩(Contextual embedding)의 유의미한 영향을 밝혀내었다.

'Idiom Principle' 을 직접적으로 적용한 연구로는 [6]이 있는데, [6]에서는 영어의 관용표현을 단일 토큰으로 취급하였다. 이를 위해 표현 가운데에 '_' 를 넣어 하나의 토큰으로 묶는 한편 묶을 때 명사, 동사 등을 레마화하여 정규화하였다. 예로 'blow the whistle' 이란 표현을 'blow_whistle' 로 만들어 실험에 사용하였다. 별도의 토큰으로 취급하는 말뭉치와 단일 토큰으로 취급하는 말뭉치를 대상으로 Word2Vec, BERT, Context2Vec로 분류 태스크를 수행하였으며, 대부분의 모델이 단일 토큰으로 취급하였을 때 0.01 내지 0.06 정도 더 좋은 f1-score를 나타냈다.

3. 실험 데이터셋

3.1. 관용표현 레이블링 말뭉치

본 연구에서는 표현이 관용적으로 쓰였는지 혹은 직설적으로 쓰였는지를 분류하는 태스크를 위한 데이터셋으로 [16]에서 소개된 '관용표현 레이블링 말뭉치' 를 사용하였다. '관용표현 레이블링 말뭉치' 는 직설적 의미와 관용적 의미 두 가지를 모두 가진 표현 15개[1]에 대해 구축되었으며, 각 표현은 '체언+격조사+용언' 형태를

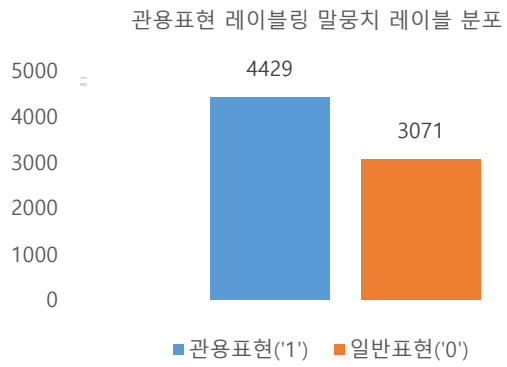
[1] Appendix 1. 참조

가진다. 말뭉치는 표현 당 500건 총 7500건으로 구축되었으며, 표현과 레이블 두 항목으로 구성된 이진 분류를 위한 말뭉치이다. 실험에서는 [그림 1]와 같이 레이블링 대상 문장인 'document'와 레이블인 'label' 항목을 입력으로 사용하였다.

	document	label
4664	거친 말투에 잠시 당황했던 원철은 남자가 고개를 저으며 문을 닫으려 하자 사력을 다...	0
4411	둔한 놈 같으니 너는 찝찝하지도 않냐 당장 눈에 띄지 않는 곳으로 보내	0
7448	내가 요즘 왜 이러지 겨울이 온 것도 아닌데 다시 긴 잠에 빠질 것 같아 자주 꿈을...	0
1919	경찰에게 돌려받은 핸드백 속에서 열쇠를 꺼낸 호경은 망설임 없이 문을 열고 집 안에...	0
1298	지금 도발하면 제재하는 이런 건데 그러면 북한이 도발하지 않는다면 북한은 그러면 그...	1
426	전투에 직접 참여하지는 않았지만 며칠 사이에 레이어드는 은닉에서 중요한 구성원으로 ...	1
6831	이와 함께 의종 년 경인년에 정중부의 무인난으로 시작된 백년간의 고려 무인집권도 막...	1
805	이런 나의 말을 듣고 당신은 영업사원이 아니므로 수치화할 수 있는 목표가 존재할 리...	0
410	호수의 크기는 중요하지 않다 갈릴리 호수가 아무리 크다 한들 내 청춘에 자리 잡고 ...	1

[그림 1] '관용표현 레이블링 말뭉치'

레이블링 말뭉치는 관용적 의미를 가진 관용표현으로 쓰였을 경우 '1', 일반적 의미를 가진 일반표현으로 쓰였을 경우 '0'으로 레이블링되었으며 [그림2]와 같은 레이블 분포를 나타냈다.



[그림 2] 관용표현 레이블링 말뭉치 레이블 분포

3.2. Idiom Principle을 적용한 데이터셋 가공

본 연구에서는 기본적으로 어절과 형태소를 토큰으로 하여 임베딩한 결과를 사전 학습에 사용하였으며, 임베딩을 위해 사전 학습에 사용된 말뭉치를 세 가지 방법으로 전처리하였다. 데이터 정교도 및 신뢰성을 높이기 위해 단순 단어 검색을 바탕으로 한 자동 전처리가 아닌 수동으로 전처리하였다.

첫 번째 임베딩은 표현 부분에 아무런 가공을 가하지 않은 임베딩이다. 즉 문맥에 나타난 표현의 표면형(surface form)을 그대로 사용하였다. 다른 하나는 'Idiom Principle'을 적용한 단순 단일 토큰 임베딩으로써 표현에 포함된 어절을 기본형 고정 없이 단순히 '_'로 연결하여 단일 토큰처럼 인식되도록 하였다. 마지막 하나는 표현을 '_'로 이은 것을 stemming하여 기본형으로 고정한 임베딩으로, stemming은 체언의 경우 격조사 복원, 용언의 경우 활용형을 모두 기본형으로 바꾸는 작업을 진행하였다. 즉 어절의 경우 '체언+격조사_용언 기본형', 형태소의 경우 '체언_격조사_용언 기본형'을 하나의 토큰으로 인식하도록 만든 stemming 단일 토큰 임베딩이다.

Train과 test를 위한 말뭉치 역시 사전 학습 말뭉치와 마찬가지로 어절과 형태소로 토크나이징 후, 세 가지로 수동 전처리하여 사용하였다([그림 3]). 이에 따라 실험에 사용된 사전 학습 임베딩 및 데이터셋은 총 6개이며, 이를 바탕으로 Word2Vec, FastText에 대해 각각 실험이 수행되었다.

	document	label
11	나 열 받게 하지 마. 안 그럼 증간 소음 확실하게 보장해줄 거니까. 사냥개도 키우...	1
12	그래 가지고 평소에 티비 보면서 갑수 선생님 보면서 열 받다가도 아 그 그때 김갑수...	1
13	만에 하나 나한테 무슨 일이 생겨도 나의 길을 간 거라고 생각하고 꿋꿋이 살아줘. ...	1
14	이런 사람은 잘 눈에 띄지 않는다. 어떻게 보면 남의 꽁무니만 쫓아가는 것처럼 보인...	0
15	누군가 자신을 부르는 소리에 도순이 고개를 돌려보니, 창가 쪽 조용한 자리에 여럿이...	0
16	벼르고 벼르던 싸움이 이번에도 싱겁게 막을 내리려는 셈인지 이때 마침 큰딸 오희(五...	1
	document	label
11	나 열_받게 하지 마. 안 그럼 증간 소음 확실하게 보장해줄 거니까. 사냥개도 키우...	1
12	그래 가지고 평소에 티비 보면서 갑수 선생님 보면서 열_받다가도 아 그 그때 김갑수...	1
13	만에 하나 나한테 무슨 일이 생겨도 나의 길을_간 거라고 생각하고 꿋꿋이 살아줘. ...	1
14	이런 사람은 잘 눈에_띄지 않는다. 어떻게 보면 남의 꽁무니만 쫓아가는 것처럼 보인...	0
15	누군가 자신을 부르는 소리에 도순이 고개를 돌려보니, 창가 쪽 조용한 자리에 여럿이...	0
16	벼르고 벼르던 싸움이 이번에도 싱겁게 막을_내리려는 셈인지 이때 마침 큰딸 오희(五...	1
	document	label
11	나 열을_받다 하지 마. 안 그럼 증간 소음 확실하게 보장해줄 거니까. 사냥개도 키...	1
12	그래 가지고 평소에 티비 보면서 갑수 선생님 보면서 열을_받다 아 그 그때 김갑수 ...	1
13	만에 하나 나한테 무슨 일이 생겨도 나의 길을_가다 거라고 생각하고 꿋꿋이 살아줘....	1
14	이런 사람은 잘 눈에_띄다 않는다. 어떻게 보면 남의 꽁무니만 쫓아가는 것처럼 보인...	0
15	누군가 자신을 부르는 소리에 도순이 고개를 돌려보니, 창가 쪽 조용한 자리에 여럿이...	0
16	벼르고 벼르던 싸움이 이번에도 싱겁게 막을_내리다 셈인지 이때 마침 큰딸 오희(五姬...	1

[그림 3] 어절 토큰 말뭉치에 대한 데이터 셋 가공 예: 표면형(위), 단일 토큰(가운데), stemming(아래)

4. 실험 및 결과

4.1. 실험 설계

3.2.에서 서술한 바와 같이 실험은 가공된 말뭉치를 이용한 사전 학습 임베딩 주입 후 분류 태스크 수행 순으로 진행하였다.

사전 학습 임베딩에는 세종 계획 문·구어 말뭉치(총 3600만 어절) 중 무작위로 선정한 11,796,245 어절을 훈련 말뭉치로 사용하였다. 사전 학습 임베딩을 위해서 어절, 형태소로 토크나이징 후 각 토크나이징 결과 별로 '표면형 말뭉치', '단순 단일 토큰 말뭉치', 그리고 'stemming 단일 토큰 말뭉치(stemming)'를 생성하였으며 Word2Vec 사전 학습 임베딩의 경우 Skip-gram으로 window 5, 300차원, 5 epochs로 학습을 진행하였다. FastText도 Word2Vec과 동일하게 window 5, 300차원, 5 epochs로 훈련을 진행하였다. 이 두 임베딩을 주입하여 수행하는 분류 태스크의 분류 레이어로는 bi-LSTM을 사용하였으며 메모리 셀 수는 512로 설정하였다. 분류 태스크 수행 시 활성화 함수로는 sigmoid 함수를 사용하였으며, 하이퍼파라미터는 배치 사이즈 16, 10 epochs, Adam optimizer, 학습률 0.001(1e-3)로 설정하였다. 훈련 중 validation은 0.2의 비율로 이루어졌다.

[표 1] Word2Vec, FastText실험에 사용된 하이퍼파라미터

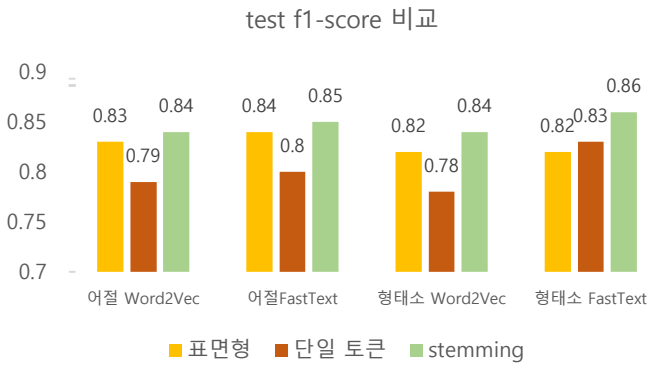
	하이퍼파라미터 설정값
Optimizer	Adam
학습률	1e-3
배치 사이즈	16
Epochs	10
Validation rate	0.2

4.2. 실험 결과

분류 태스크 수행은 4.1.에서 설명한 임베딩에 상응하는 가공 데이터셋을 사용하여 수행되었다. 수행 결과 전체 분류 성능은 다음 표와 같이 나타났다.

[표 2] 실험 test 분류 성능

어절	Test 성능	표면형	단일 토큰	Stemming
Word2Vec	accuracy	0.8	0.75	**0.81**
	presicion	0.8	0.77	0.83
	recall	0.88	0.82	0.85
	f1-score	0.83	0.79	**0.84**
FastText	accuracy	**0.82**	0.75	**0.82**
	presicion	0.78	0.74	0.83
	recall	0.90	0.88	0.87
	f1-score	0.84	0.8	**0.85**
형태소	**Test 성능**	**표면형**	**단일 토큰**	**Stemming**
Word2Vec	accuracy	0.81	0.74	**0.82**
	presicion	0.86	0.75	0.86
	recall	0.81	0.83	0.82
	f1-score	0.83	0.78	**0.84**
FastText	accuracy	0.8	0.8	**0.84**
	presicion	0.84	0.82	0.87
	recall	0.81	0.86	0.85
	f1-score	0.82	0.83	**0.86**



[그림 4] test f1-score 비교

실험 결과, Word2Vec과 FastText 모두 어절, 형태소 토큰 실험에서 단순 단일 토큰의 성능이 가장 낮았으며, 전반적으로 stemming을 한 단일 토큰의 accuracy 및 f1-score가 표면형, 단순 단일 토큰 임베딩에 비해 높게 나타났다. 또한 어절 토큰보다 형태소 토큰을 기반으로 실험을 진행했을 때 분류 성능이 향상되었으며, Word2Vec보다 FastText의 분류 성능이 좋았다. [1]에서도 형태론적 특성을 띈 프랑스 어, 바스크 어에 대해 Multi Word Expression 분류 task를 Word2Vec과 FastText에 대해 수행하였는데, FastText의 분류 성능이 Word2Vec에 비해 앞섰다. 이는 형태론적 특징이 강하게 드러나는 언어에 대해 Word2Vec보다 FastText가 강건한 성능을 보이는 것으로 볼 수 있다. 어절 토큰보다 형

태소 토큰의 분류 성능이 앞서는 것으로 미루어 보아 한국어 역시 교착어에 속하기 때문에 형태론적 특성을 강하게 지니는 것을 알 수 있으며, 또한 형태론적 특성에 강건함을 보이는 FastText의 성능이 Word2Vec에 앞섰다.

가장 높은 accuracy와 f1-score를 기록한 실험은 형태소 stemming 임베딩을 사용한 FastText로 accuracy 0.84, f1-score 0.86을 기록하였다. 같은 단일 토큰이라도 용언을 stemming하여 단일 토큰화한 임베딩이 곡용·활용형을 고정하지 않았던 단순 단일 토큰의 임베딩보다 분류 성능이 증가하는 것을 관찰할 수 있는데, 이는 한국어의 교착어적 특성으로 인해 발생하는 현상으로 볼 수 있다. 단일 토큰의 경우 체언의 곡용형이나 용언의 활용형이 고정되지 않기 때문에 표현에 대한 매 토큰들이 서로 다른 벡터로 취급되기 때문이다. 곧 표현 '말을 듣다'의 곡용·활용형인 '말을_듣고(어절 토큰)', '말_을_듣고(형태소 토큰)', '말을_듣다', '말_을_듣다', '말을_들어서', '말_을_들어서', '말_들어'가 모두 다른 토큰으로 취급됨으로써 분류 성능에 부정적인 영향을 끼치게 된다. 반면 stemming을 진행한 경우 위의 곡용·활용형이 모두 '말을_듣다(어절 토큰)' 혹은 '말_을_듣다(형태소 토큰)'란 기본형으로 고정되기 때문에 기본형의 벡터와 주변 문맥 간 벡터의 상관관계가 비교적 명확히 표상될 수 있다.

Word2Vec, FastText의 결과는 관용표현 중의성 해소에 있어 관용표현이 단일 토큰으로 취급될 필요성을 시사한다. 특히 두 실험의 stemming 결과로 미루어 보아, 관용표현과 같은 단어 이상의 표현들을 인간의 언어적 직관과 마찬가지로 단일 토큰으로 처리하는 것은 어느 정도 의미가 있는 것으로 유추된다.

5. 결론 및 향후 연구

본 연구에서는 표현의 일반표현 및 관용표현 의미 간 중의성 해소를 위해 'Idiom Principle'을 적용하여 어절 단위, 형태소 단위 말뭉치에 대해 다양한 층위(표면형, 단일 토큰, stemming)의 임베딩을 생성하여 사전 학습 임베딩 및 분류 태스크 실험을 수행하였다. 즉 사전 학습된 임베딩과 분류 실험을 진행할 때 어절 토큰, 형태소 토큰 별 말뭉치를 표면형, 단순 단일 토큰, stemming 단일 토큰 세 가지로 가공하여 사용하였다.

실험 수행 결과, 전반적으로 어절 토큰보다 형태소 토큰을 기반으로 했을 때 분류 성능이 높았으며, 또 FastText의 분류 성능이Word2Vec보다 높았다. 이로 미루어 보아 한국어를 자연어 처리할 때 형태소를 기반으로 실험하는 것이 필수적임을 확인할 수 있었다. 또한 대부분의 실험에서 'Idiom Principle'을 반영한 stemming 단일 토큰을 적용했을 때의 accuracy 및 f1-score가 다른 경우들에 비해 높았다. 이를 토대로 언어 사용자의 인식과 같이 자연어 처리 과정에서도 'Idiom Principle'이 어느 정도 영향을 미치는 것을 확인할 수 있었으며, 관용표현과 같은 비합성성(non-compositionality)을 가진 표현들이 임베딩 혹은 훈련 과정 중 어떻게 처리되는지에 대한 실마리를 얻을 수 있었다.

다만 본 연구에서는 단어 기반 임베딩인 Word2Vec과 FastText만을 사용하였으나, Contextual representation을 기반으로 하는 ELMo, Context2Vec 혹은 BERT에 대해서는 실험을 진행하지 않았다. 이에 향후 연구로는 문맥을 기반으로 한 모델들에도 'Idiom Principle'이 적용될 수 있는지 추가적인 연구를 진행하고자 한다. 또한 'Idiom Principle'을 적용한 결과의 향상이 단순히 토큰 종수가 줄어 성능이 향상된 것인 지, 혹은 task 수행 과정 중 통사·의미론적 정보의 유의미한 작용인지에 대한 추가 연구 역시 필요하다.

참고문헌

[1] Nicolas Z ampieri, C arlos Ramish , Geraldine Damnati, Th e Impact of W ord R epresentations on Sequential Neural MWE Identification, Proceedings of the Joint Workshop on Multiword Expressions and WordNet, p.169-175, 2019.

[2] Rafael Ehren et al., Supervised Disambiguation of German Verbal Idioms with a BiLSTM Architecture, Proceedings of the Second Workshop on Figurative Language Processing, pp.211-220, 2020.

[3] Jing Peng and Anna Feldman, Experiments in Idiom Recognition, Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: Technical Papers, pp.2752-2761, 2016.

[4] 이희자, 현대 국어 관용구의 결합 관계 고찰, 대동문화연구 30호, 1995.

[5] Sinclair, J., Corpus, Concordance, Collocation.Oxford University Press,1991.

[6] Reyhaneh hashempour, Aline Villavicencio, Leveraging Contextual Embeddings and Idiom Principle for Detecting Idiomaticity in Potentially Idiomatic Expressions Proceedings of the Workshop on Cognitive Aspects of the Lexicon, pages 72-80, 2020.

[7] 문금현, 국어의 관용표현 연구, 서울대학교 대학원 박사학위 논문, 1996.

[8] 김진해, 관용어의 통사·의미론적 제약 연구, 경희대학교 대학원 석사학위 논문, 1995.
[9] Afsaneh Fazly, Suzanne Stevenson, Automatically Constructing a Lexicon of Verb Phrase Idiomatic Combinations, 11st Conference of the European Chapter of the Association for Computational Linguistics, Proceedings of the Conference, pp.337-344, 2006.
[10] Afsaneh Fazly, Paul Cook, Suzanne Stevenson, Unsupervised Type and Token Identification of Idiomatic Expressions, Computational Linguistics, Vol 35, pp.61-103, 2009.
[11] Caroline Sporleder and Linlin Li, Unsupervised Recognition of Literal and Non-Literal Use of Idiomatic Expressions, Proceedings of the 12th Conference of the European Chapter of the ACL, pp.754-762, 2009.
[12] Hessel Haagsma, Malvina Nissim, Johan Bos, The Other Side of the Coin: Unsupervised Disambiguation of Potentially Idiomatic Expressions by Contrasting Senses, Proceedings of the Joint Workshop on Linguistic Annotation, Multiword Expressions and Constructions, pp.178-184, 2018
[13] 김한샘, 한국어 숙어 연구, 한국문화사, 2011.
[14] 문경희, 이종혁, 김정인, 양기주, 일-한 기계 번역 시스템: 연어 패턴을 이용한 어휘 다의성 해소, 정보과학회논문지(B)25(8), p.1270~1280, 1998
[15] Kurfali, Murathan, and Robert Östling. "Disambiguation of Potentially Idiomatic Expressions with Contextual Embeddings." Joint Workshop on Multiword Expressions and Electronic Lexicons, 2020.
[16] 박서윤, 사전 훈련 언어 모델을 활용한 관용표현 분류 연구, 국내석사학위논문 연세대학교 대학원, 2021.

Appendix

1. 실험에 사용된 관용표현 목록[16]

자리를 잡다
-문자적 의미: 사람이 앉을 수 있는 설비나 지정 장소를 점유하다.
-관용적 의미: 「1」일정한 지위나 공간을 차지하다
「2」생각이 마음속에 뿌리를 박은 듯 계속 남아 있다.

말을 듣다
-문자적 의미: 사람의 생각이나 느낌 따위를 목구멍을 통하여 조직적으로 나타내는 소리를 듣다.
-관용적 의미: 「1」꾸지람을 듣거나 시비의 대상이 되다
「2」기계, 도구 따위가 다루는 사람의 뜻대로 움직이다.

말이 되다
-문자적 의미: 단어, 구, 문장 따위가 성립하다.
-관용적 의미: 「1」말하는 것이 이치에 맞다.
「2」어떤 사실에 대하여 서로 간에 말이 이루어지다.

문을 열다
-문자적 의미: 드나들거나 물건을 넣었다 꺼냈다 하기 위하여 틔워 놓은 곳. 또는 그곳에 달아 놓고 여닫게 만든 시설을 열다.
-관용적 의미: 「1」영업 따위를 시작하다.
「2」문호를 개방하다.
「3」어떤 조직에서 장벽을 두지 아니하고 사람을 받아들이다.
「4」어떤 일이 장애로 막히거나 중단되었던 것을 터놓다.

가슴이 아프다
-문자적 의미: 배와 목 사이의 앞부분에 통증이 생기다.
-관용적 의미: 슬프거나 안타까운 감정이 들다.

열을 받다
-문자적 의미: 물체가 뜨거워지다.
-관용적 의미: 흥분되다, 화나다.

길을 가다
-문자적 의미: 사람이나 동물 또는 자동차 따위가 지나갈 수 있게 땅 위에 낸 일정한 너비의 공간을 통해 이동하다.
-관용적 의미: 「1」사람이 삶을 살아가거나 사회가 발전해 가는 데에 지향하는 방향, 지침, 목적이나 전문분야를 지향하다.

길을 걷다
-문자적 의미: 사람이나 동물 또는 자동차 따위가 지나갈 수 있게 땅 위에 낸 일정한 너비의 공간을 다니다.
-관용적 의미: 「1」사람이 삶을 살아가거나 사회가 발전해 가는 데에 지향하는 방향, 지침, 목적이나 전문분야를 수행하다.
「2」어떤 자격이나 신분으로서 주어진 도리나 임무를 수행하다.

눈에 띄다
-문자적 의미: 물체의 존재나 형상이 보이다.
-관용적 의미: 두드러지게 드러나다.

문을 닫다
-문자적 의미: 드나들거나 물건을 넣었다 꺼냈다 하기 위하여 틔워 놓은 곳. 또는 그곳에 달아 놓고 여닫게 만든 시설을 막다.
-관용적 의미: 경영하던 일을 그만두고 폐업하다.

못을 박다
-문자적 의미: 목재 따위의 접합이나 고정에 쓰는 물건을 두들겨 쳐서 꽂히게 하다.
-관용적 의미: 「1」다른 사람에게 원통한 생각을 마음속 깊이 맺히게 하다.
「2」어떤 사실을 꼭 집어 분명하게 하다.

뿌리를 내리다
-문자적 의미: 식물의 밑동이 땅속으로 들어가다.
-관용적 의미: 정착하다, 자리를 잡다.

몸에 배다
-문자적 의미: 신체에 냄새가 스며들어 오래도록 남아 있다.
-관용적 의미: 여러 번 겪거나 치러서 아주 익숙해지다.

막을 내리다
-문자적 의미: 무대 앞을 가리는 천을 밑으로 늘어뜨리다.
-관용적 의미: 무대의 공연이나 어떤 행사를 마치다.

꿈을 꾸다
-문자적 의미: 잠자는 동안에 깨어 있을 때와 마찬가지로 여러 가지 사물을 보고 듣는 정신 현상이 발생하다.
-관용적 의미: 희망하다.