Journal of the Korea Convergence Society
Vol. 10. No. 6, pp. 41-47, 2019

ISSN 2233-4890
https://doi.org/10.15207/JKCS.2019.10.6.041

조세심판 문서 검색 효율 향상 모델에 관한 연구

이후영[1], 박구락[2*], 김동현[3]
[1]공주대학교 컴퓨터공학과 박사과정, [2]공주대학교 컴퓨터공학부 교수, [3]공주대학교 컴퓨터공학과 박사

A Study on the Improvement Model of Document Retrieval Efficiency of Tax Judgment

Hoo-Young Lee[1], Koo-Rack Park[2*], Dong-Hyun Kim[3]
[1]Ph.D. Student, Dept. of Computer Engineering, Kongju National University
[2]Professor, Dept. of Computer Science & Engineering, Kongju National University
[3]Ph.D., Dept. of Computer Engineering, Kongju National University

요 약 조세 심판에 대한 선결정례는 법원 판례의 경우 유사 심판례를 검색하여 파악하는 것이 매우 중요한 상황이다. 그러나 기존 심판문에 대한 검색은 사용자가 입력하는 키워드를 통하여 검색하는 방법을 사용하고 있으나, 정확한 키워드의 입력이 필요하며, 키워드를 모르는 경우 필요한 문서를 검색하는 것은 불가능하다. 또한 검색된 문서 중에는 내용이 다른 경우도 발생한다. 이에 본 논문에서는 정확한 심판례의 검색을 위하여 문서를 3차원 공간에 벡터화하고, 코사인 유사도를 계산하여, 거리상 가까운 문서를 검색하는 방법의 효율성을 향상시키기 위하여 심판례에서 사용되고 있는 단어들의 유사도를 분석한 후, 최빈값을 추출하여 본문의 텍스트에 삽입하는 방법으로 검색하고자 하는 문서의 코사인 유사도를 향상시키는 방안을 제안한다. 제안 모델을 통하여 조세와 관련된 심판례를 검색하고자 하는 사용자에게 신속하고, 정확한 검색을 제공할 수 있을 것으로 기대된다.

주제어 : 융합, 조세 판례, 유사 문서, 자연어처리, 워드 임베딩

Abstract It is very important to search for and obtain an example of a similar judgment in case of court judgment. The existing judge's document search uses a method of searching through key-words entered by the user. However, if it is necessary to input an accurate keyword and the keyword is unknown, it is impossible to search for the necessary document. In addition, the detected document may have different contents. In this paper, we want to improve the effectiveness of the method of vectorizing a document into a three-dimensional space, calculating cosine similarity, and searching close documents in order to search an accurate judge's example. Therefore, after analyzing the similarity of words used in the judge's example, a method is provided for extracting the mode and inserting it into the text of the text, thereby providing a method for improving the cosine similarity of the document to be retrieved. It is hoped that users will be able to provide a fast, accurate search trying to find an example of a tax-related judge through the proposed model.

Key Words : Convergence, Tax Cases, Similar Documents, NLP, Word Embedding

*Corresponding Author : Koo-Rack Park(ecgrpark@kongju.ac.kr)
Received May 1, 2019 Revised June 3, 2019
Accepted June 20, 2019 Published June 28, 2019

1. 서론

데이터의 중요성에 따라 대량의 문서 데이터가 증가하고 있으나, 일반 사용자는 해당하는 문서를 이해하기 위하여 요약된 정보를 필요로 하고 있고, 이러한 정보는 단순한 통계에 의해 생성되고 있어, 의미 있는 문장의 생성과 문장의 모호성을 해결하기 위한 연구가 미흡한 실정이다[1]. 문서 데이터에 대한 연구는 학술연구정보서비스 사이트의 데이터베이스를 활용하여 키워드 추출과 키워드 사이의 연결 중심성을 시각화하여 연구 동향 분석방법을 제시하는 연구[2], 웹 추천기법에서 사용하기 위하여 Word2Vec 기법과 앙상블 합성곱 신경망을 이용하여 영화 추천 방안에 대한 연구[3]와 같이 다양하게 이루어지고 있다. 또한 4차 산업혁명의 대표적인 키워드인 인공지능을 산업의 여러 분야에 융합하고자하는 연구가 활발하게 진행되고 있으며, 특히 스마트 팩토리 효율성 강화 모델에 관한 연구와 보건의료분야에서의 인공지능 기술에 대한 연구와 같이 다양한 분야에서 인공지능을 활용한 연구가 이루어지고 있는 실정이다[4,5]. 정보기술을 이용하여 많은 분야의 기술들이 자체적으로 융합되어 새로운 형태의 기술과 제품들이 개발되고 있는 상황에서, 최근 인공지능을 활용한 법령해석과 적용에 대한 노력은 인공지능에 대한 관심이 높아지면서 연구가 활발하게 진행되고 있다. 특히 조세 심판에 대한 선결정례가 중요한 법원의 판례의 경우 사건의 유사 심판례를 찾는 노력은 더욱 중요하다. 기존의 심판문 검색은 사용자가 입력하는 키워드를 통해서 문서를 검색하는 방법을 사용하고 있다. 키워드 검색은 문서 내에서 등장하는 키워드를 찾아 검색 결과를 표시하기 때문에 정확한 키워드를 입력해야 하며, 정확한 키워드를 모르는 경우 원하는 문서를 찾을 수 없는 단점을 가지고 있다. 해당 키워드를 포함하지만 내용이 전혀 다른 경우도 검색되어 사용자가 정확한 심판례를 검색하는데 많은 불편함을 호소하고 있는 상황이다. 본 논문에서는 정확한 심판례를 검색하기 위하여, 심판례 문서를 3차원 공간에 벡터화하고, 코사인 유사도를 계산하여 거리상 가까운 문서를 검색하는 방법에 대한 효율성을 향상시키기 위하여, 심판례에서 사용하는 단어들의 유사도를 분석한 후, 최빈값을 추출하여 이를 본문의 텍스트에 삽입하는 방법으로, 문서의 코사인 유사도를 향상시키는 방법을 제안한다. 제안 모델을 통하여 조세와 관련된 심판례를 검색하고자 하는 사용자에게 정확한 심판례의 검색이 편리할 것으로 기대된다.

2. 관련연구

2.1 워드 임베딩

워드 임베딩은 문장과 문단을 포함하는 각 단어 사이의 관계를 수치화하여 벡터로 표현하는 기법이다. 이는 특정한 단어의 의미와 주변 단어 분포가 관계성이 있음을 가정하며 각 단어 벡터 값의 연산을 통하여 단어 간 관계를 만든다[6,7]. 심층 신경망을 이용한 단어 임베딩 기반 학습은 NNLM(neural network language model)이 제안된[8] 이후 분류에 활발하게 적용되고 있다[9,10].

Word2Vec은 단어 임베딩을 활용한 텍스트 및 이미지 분류, 자연어 처리, 질의응답 시스템 연구에 폭넓게 활용되고 있다. Word2Vec은 은닉층의 수를 줄임으로 모델을 단순화시켜 학습 시간을 단축하고 단어 간의 관계를 추론할 수 있다[11].

단어의 의미와 순서를 가지고 있는 벡터의 형태로 단어를 표현하는 기법인 Word2Vec 모델은 특정한 임베딩 공간상에서 같은 맥락을 가지고 있는 단어들이 가까운 거리를 가지고 있다는 것을 전제로 출발하고 있으며[12,13], 주어진 문장을 구성하는 단어들의 전후 관계를 학습하여 단어의 의미를 내포하고 있는 벡터 값으로 문서를 구성하고 있는 자질들을 수치화한다. 이것은 기존의 통계적인 방식을 활용한 연구[14-16]와는 다르게 별도의 유사도 계산이나 차원을 축소하는 과정 없이 변별적인 특징을 내포하고 있는 벡터 값으로 단어를 수치화할 수 있다.

2.2 Doc2Vec

Doc2Vec은 비정형 데이터인 문장이나 혹은 구문, 문단을 벡터 공간에서 표현하는 알고리즘이며 패러그래프 벡터와 동일한 개념이다[17]. Word2Vec이 단어를 임베딩하기 위한 모델로 제안되었듯이, Doc2Vec은 문서 임베딩 모델을 활용하여 벡터 공간 안에 표현하는 기법으로 문장이나 문서에 대하여 벡터 값으로 나타내는 비지도 학습 방법이다[18].

Doc2Vec 모델은 두 가지 알고리즘인 Distributed Memory(DM)와과 Distributed Bag of Words(DBOW) 방식으로 구분할 수 있다. DM모델은 다음의 Fig. 1과 같이 단어를 학습시킬 경우 각각의 학습 단계를 벡터에 기억하게 하고, 이후 최종 학습된 벡터를 정의하는 방식이다[19].

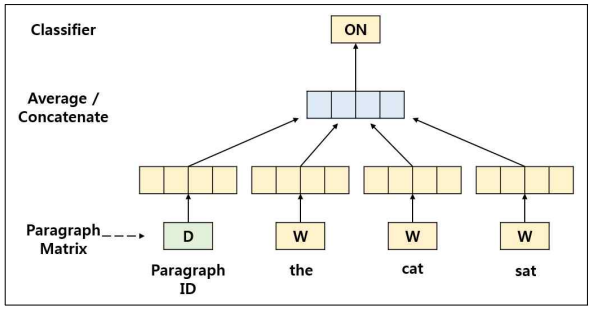

Fig. 1. Distributed Momory(DM) Model

DBOW 모델은 다음의 Fig. 2와 같이 어떠한 문서가 주어졌을 경우, 그 문서에 포함되어 있는 단어를 예측할 수 있는 알고리즘이다[19].

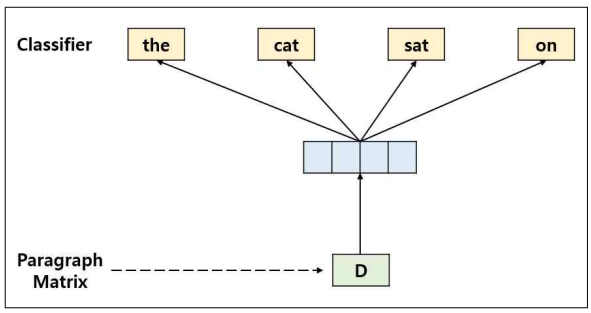

Fig. 2. Distributed Bag of Words(DBOW) Model

2.1 NLP(Natural Language Processing)

다음의 Fig. 3은 자연어를 처리하는 과정을 단계별로 나타낸 것이다[20].

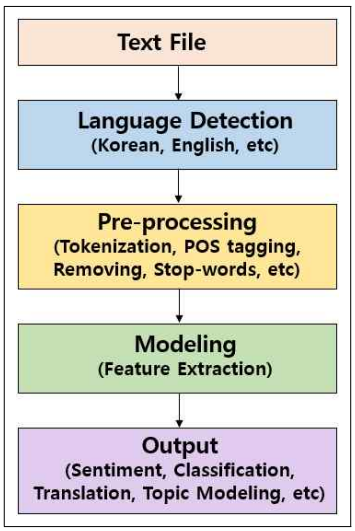

Fig. 3. Flow of Natural Language Processing

자연어처리(NLP)는 인간이 발화하는 언어인 자연어의 현상을 컴퓨터가 이해할 수 있는 형태로 분석하여 처리하고, 다시 이해할 수 있는 자연어적 형태로 표현하는 기술을 의미하고 있다. 자연어처리 방식으로는 전통적으로 통계기반과 규칙기반의 접근법이 있고, 이 둘의 장점을 통합한 하이브리드 방식이 있으며, 딥러닝과 같은 인공신경망 방식으로 분류할 수 있고, 딥러닝 방식은 입력 및 출력 문장을 하나의 쌍으로 하여, 가장 적합한 번역 및 표현의 결과를 찾는 방식으로 최근 들어 주로 활용하고 있다[21].

3. 제안 모델

3.1 시스템 구성도

제안 모델은 문장을 벡터로 표현하는 기법인 Doc2Vec 기법을 활용하여 문서의 내용을 분석하여 벡터화 하고 이를 통해 특정 문서와 가까운 거리에 있는 문서를 찾아내는 기존의 방법에서 더 나아가 유사도의 향상을 위하여 Word2Vec을 활용하여 유사한 단어들의 집합을 찾아내고 이를 문서에 삽입함으로 문서의 유사도를 높이는 방법을 제안한 것으로 다음의 Fig. 4는 제안 모델의 시스템 구성도이다.

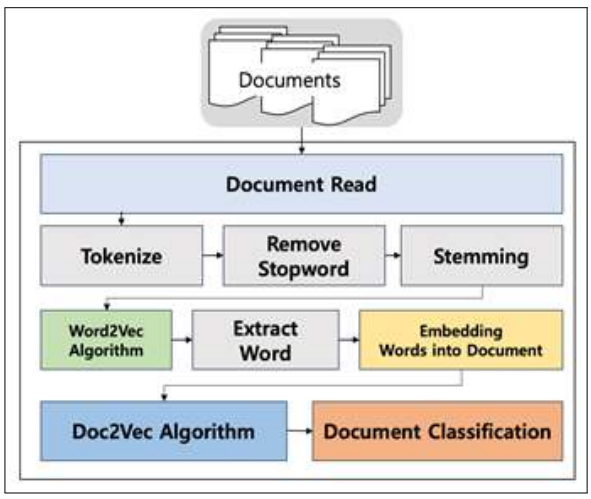

Fig. 4. System Configuration

분석하고자 하는 심판례 문서를 수집한 후, 수집된 문서를 분석하여 불필요한 문단들을 제거하고, 분석에 사용될 주요 문단들을 추출하여 이를 정제하는 과정을 진행한다. 정제 과정은 단어들을 어절 단위로 분리(Tokenizing)

하는 작업을 수행하고, 분리된 단어 중에서 분석에 불필요한 단어들을 제거하는 작업을 수행하며, 형태론 및 정보 검색 분야에서 어형이 변형된 단어로부터 접사 등을 제거하고, 단어의 어간을 분리해 내는 어간 추출을 수행한다. 추출한 단어 중에서 명사이며, 최소 두 글자 이상인 단어와 10회 이상 등장하는 단어로만 사전을 구축하였다. 다음의 Table 1은 상위 10개의 단어만 추출한 표로서, '처분' 라는 단어는 43,268회 등장하여 가장 많이 등장하였고 다음으로 청구, 청구인, 쟁점, 법인 등의 순으로 빈도가 추출되었다.

Table 1. Extract the Top 10 Words

No.	Word	Count
1	Disposition	43,268
2	Claim	42,302
3	Claimant	33,373
4	Issue	29,774
5	Corporation	25,883
6	Taxation	21,338
7	Request	15,723
8	Relevant	14,043
9	Mistake	13,962
10	Acquisition	12,417

3.2 문서 구조 분석

다음의 Fig. 5는 심판문의 구조이다. 심판문은 비정형 데이터로 분류할 수 있으나 몇 가지 정형화된 요소를 가지고 있다.

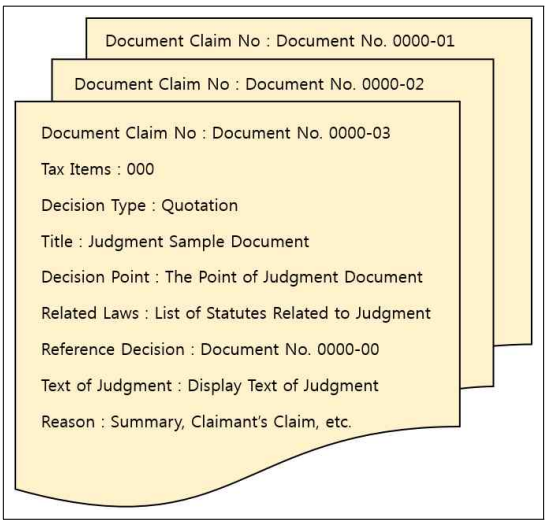

Fig. 5. Structure of Judgment Document

심판문은 문서번호, 심판문의 특징을 가장 잘 나타내는 세목과 심판문의 요약본이라고 할 수 있는 결정요지와 심판문 전체의 내용에 해당하는 이유 등으로 구성되어 있으며, 심판 제목과 결정요지, 주문에 있는 내용을 추출하여 분석하였다.

4. 실험 및 고찰

본 논문에서 사용한 심판판례는 조세에 관련된 심판을 수행하는 국무조정실 산하기관인 조세심판원에서 검색 가능한 2014년~2018년까지의 5년간의 심판례를 수집하여 분석하였다. 심판례 문서는 반정형 문서로 제목, 결정요지, 관련법령, 주문, 이유, 결론 등의 문서 형식을 가지고 있다. 심판례 문서의 문서번호를 레이블로 하고 제목과 결정요지, 관련 법령을 데이터로 분류한 후, 문서번호를 제외한 텍스트 문서를 한글 자연어처리에서 일반적으로 사용하고 있는 KoNLPy를 통하여 형태소를 분석하였다.

두 글자 이상 명사를 추출하고, 일부 불필요한 특수문자들을 제거한 후 Word2Vec 기법을 활용하여 각 단어들을 벡터 형태로 나타내었고, 분석의 용이성을 위해 10회 이상 반복되는 단어만 활용하였다. 다음의 Table 2는 최빈 단어와 유사도가 가장 높은 단어의 리스트 일부를 정리한 것이다.

Table 2. Most Words and Highly Similar Word List

No	Word	Count	Similarity
1	Check	6,648	Tax(0.9222)
2	Expense	1,882	Necessity(0.8545)
3	Bill	790	Delivery(0.8359), Payment- of-Taxes(0.7233)
4	Issue-Farmland	775	Cultivation(0.7112)
5	Agreement	697	Country-of-Origin(0.8235), Duty(0.7045), Asean(0.7032)
6	Delivery	600	Bill(0.8359)
7	Constitution	531	Constitutional-Court(0.8540), Equality0.7012)
8	Constitutional-Court	525	Constitution(0.8540), Provision(0.7416), Unconstitutional(0.7106)
9	Corruption	519	Fraud(0.8180)
10	Country-of-Origin	490	Agreement(0.8235), Asean(0.7244)

단어 간의 거리를 추출하기 위해서 사용한 패키지는 NLP에서 주제 모델링, 문서 인덱싱 및 대형 코퍼스와의 유사성 검색하는데 높은 성능을 발휘하는 파이썬 라이브러리인 Gensim을 사용하였다. Table 2는 유사단어 테이블로 출현 빈도수가 높은 n개의 단어에 대하여 유사어를 추출한 것으로, 단어 간의 벡터 거리가 짧아도 문맥상 의미가 부족한 단어들은 임의로 제거하였으며, 10회 이상 등장하는 단어와 유사도가 1에 근접할수록 높다고 할 때 0.7 이상의 단어들을 사용하였다. 그 결과 10,469의 워드리스트에서 520개의 워드리스트로 데이터 테이블을 생성하였다.

또한 유사도는 1에 근접할수록 높은 유사도를 나타내고 있다. 계산서는 문서에서 6,648회 등장하며 세금(0.9222)과 높은 유사도를 보이고, 고지서는 문서에서 790회 등장하며 송달(0.8235), 납세(0.7233) 등의 단어와 높은 유사도를 보인다. 그러나 일부 단어들은 단어의 의미가 왜곡되어 있는 경우와 형태소 분석의 오류로 인한 단어들은 제외하였다. 최빈 단어와 높은 유사도를 가지는 단어를 해당 문장에 함께 삽입한 후 Doc2Vec기법을 활용하여 문서 제목으로 문서간의 거리를 계산하고, 전후 관계를 비교하였다. 단, Doc2Vec의 분석환경 변수는 변화를 주지 않고, 단순히 유사 단어를 삽입한 것과 삽입하지 않은 것에 대한 실험을 진행하였다.

다음의 Fig. 6은 텐서플로우에서 제공하는 고차원 데이터의 시각화 툴로 벡터 데이터를 3차원으로 시각화한 것이며, 기준 문서인 조심2015지0619를 기준으로 각 문서의 거리를 표시하였다. 포인트는 3차원 공간의 원근에 따라 크기에 차이가 있으나 실제로는 동일한 크기이다.

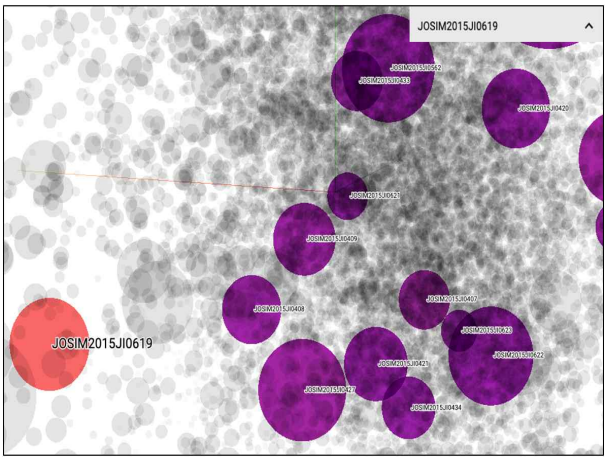

FIg. 6. Visualization of Analytical Data

다음의 Table 3은 일부 문서의 코사인 유사도의 변화를 나타낸 것이다. 실험결과와 같이 문서에 유사단어를 입력한 데이터가 문서 간의 유사도에 영향을 미치는 것을 확인하였다.

Table 3. Change in Cosine Similarity

Standard Document	Similar Document	Cosine Similarity 1	Cosine Similarity 2	Variation
JOSIM2018SE4254	JOSIM2018SE4263	0.027	0.015	0.012(−)
	JOSIM2018SE4113	0.125	0.114	0.011(−)
JOSIM2013JI0740	JOSIM2014JI0054	0.052	0.047	0.005(−)
	JOSIM2013JI0976	0.052	0.043	0.009(−)
JOSIM2015JI0619	JOSIM2015JI0622	0.045	0.041	0.004(−)
	JOSIM2015JI0621	0.047	0.040	0.007(−)

기준 문서인 조심2018서4254를 기존의 Doc2Vec의 기법으로 문서를 분석했을 경우, 조심2018서4263, 조심2018서4113의 문서가 거리상 가까운 문서로 표시되었고, 코사인 유사도 1은 각각 0.027, 0.125로 표시되었다. 다음으로 유사 단어를 입력하고 같은 조건으로 분석한 경우 조심2018서4263 문서는 코사인 유사도 2와 같이 0.015로 기준문서와의 거리가 0.012 감소하였고, 조심2018서4113 문서는 0.114로 거리가 0.011로 감소하여 기준 문서와의 거리가 가까워졌음을 확인하였다.

5. 결론 및 향후 연구 방향

제안 모델의 실험 결과 심판문을 분석하여 유사단어를 추출하고, 해당 유사단어들을 심판문에 삽입하는 방법으로 벡터 공간에서 각 문서들의 거리에 영향을 미치는 것을 확인하였다. 본 논문에서 가장 중요한 선행 작업은 유사단어를 어떻게 추출하고 어떻게 적용할 것인가에 대한 정의이다. 이에 따라 적절한 단어 군을 삽입한 문서들은 그 유사도가 줄어들고, 그렇지 않은 경우는 유사도가 증가하는 것을 확인할 수 있었다. 향후 연구에서는 양질의 유사 단어의 집합을 만드는 작업을 위해 최대한 많은 양의 심판문을 분석하는 것이 필요하고, 적절한 제약조건을 통해 집합의 질을 높여가는 연구가 필요하다. 이와 병행하여 기사 검색, 도서 검색, 특허 혹은 논문 등 의미론적 검색이 가능한 다양한 분야에 적용할 수 연구가 계속되어야 할 것이다.

REFERENCES

[1] S. J. Baek. (2017). Multi-Document Summarization Method Based on Semantic Relationship using VAE. *Journal of Digital Convergence, 15(12),* 341-347.
DOI : 10.14400/JDC.2017.15.12.341

[2] W. J. Lee & T. G. Kim. (2019). A Study on the Research Trend in the Dyslexia and Learning Disability Trough a Keyword Network Analysis. *Journal of Digital Convergence, 17(1),* 91-98.
DOI : 10.14400/JDC.2019.17.1.091

[3] B. S. Kang. (2019). A Study on the Accuracy Improvement of Movie Recommender System Using Word2Vec and Ensemble Convolutional Neural Networks. *Journal of Digital Convergence, 17(1),* 123-130.
DOI : 10.14400/JDC.2019.17.1.123

[4] Y. S. Jeong. (2019). A Model Design for Enhancing the Efficiency of Smart Factory for Small and Medium-Sized Businesses Based on Artificial Intelligence. *Journal of Convergence for Information Technology, 9(3),* 16-21.
DOI : 10.22156/CS4SMB.2019.9.3.016

[5] J. M. Kim. (2017). Study on Intention and Attitude of Using Artificial Intelligence Technology in Healthcare. *Journal of Convergence for Information Technology, 7(4),* 53-60.
DOI : 10.22156/CS4SMB.2017.7.4.053

[6] J. Turian, L. Ratinov & Y. Bengio. (2010). Word representations: a simple and general method for semi-supervised learning. *In Proceedings of the 48th annual meeting of the association for computational linguistics,* 384- 394.

[7] T. Mikolov, I. Sutskever, K. Chen, G. S. Corrado & J. Dean. (2013). Distributed representations of words and phrases and their compositionality. *In Advances in neural information processing systems,* 3111-3119.

[8] Y. Bengio, R. Ducharme, P. Vincent & C. Jauvin. (2003). A neural probabilistic language model. *Journal of Machine Learning Research, 3(Feb),* 1137-1155.

[9] D. Tang, B. Qin & T. Liu. (2015). Document modeling with gated recurrent neural network for sentiment classification. *In Proceedings of the 2015 conference on empirical methods in natural language processing,* 1422-1432.
DOI : 10.18653/v1/d15-1167

[10] R. Collobert & J. Weston. (2008, July). A unified architecture for natural language processing: Deep neural networks with multitask learning. *In Proceedings of the 25th International Conference on Machine Learning, ACM,* 160-167.
DOI : 10.1145/1390156.1390177

[11] J. H. Yuk & S. Min. (2018). A Study of Research on Methods of Automated Biomedical Document Classification using Topic Modeling and Deep Learning. *Journal of Korean Society for Information Society, 35(2),* 63-88.
DOI : 10.3743/KOSIM.2018.35.2.063

[12] I. S. Kang. (2013). A Comparative Study on Using SentiWordNet for English Twitter Sentiment Analysis. *Journal of Korean Institute of Intelligent Systems, 23(4),* 317-324.
DOI : 10.5391/JKIIS.2013.23.4.317

[13] T. Mikolov, K. Chen, G. Corrado & J. Dean, (2013). Efficient estimation of word representations in vector space. *arXiv preprint arXiv:1301.3781.*

[14] M. Y. Ren & S. Kang. (2015). Comparison Between Optimal Features of Korean and Chinese for Text Classification. J*ournal of Korean Institute of Intelligent Systems, 25(4),* 386-391.
DOI : 10.5391/JKIIS.2015.25.4.386

[15] D. W. Lee, S. H. Baek, M. J. Park, J. H. Park, H. W. Jung & J. H. Lee. (2012). Document Summarization Using Mutual Recommendation with LSA and Sense Analysis. *Journal of Korean Institute of Intelligent Systems, 22(5),* 656-662.
DOI : 10.5391/JKIIS.2012.22.5.656

[16] S. H. Jun. (2015). A big data preprocessing using statistical text mining. *Journal of Korean Institute of Intelligent Systems, 25(5),* 470-476.
DOI : 10.5391/JKKIS.2015.25.5.470

[17] Q. Le & T. Mikolov. (2014). Distributed representations of sentences and documents. *In Proceedings of the 31st International Conference on Machine Learning (ICML 2014),Beijing, China,* 1188-1196.

[18] J. H. Lau & T. Baldwin. (2016). An empirical evaluation of doc2vec with practical insights into document embedding generation. *arXiv preprint arXiv:1607.05368.*

[19] J. H. Kim. (2019). Method of Keyword Recommendation Considering Importance and Correlation of Words. *Chosun University, Master's Thesis.*

[20] A. R. Song & Y. H. Park. (2018). WV-BTM: A Technique on Improving Accuracy of Topic Model for Short Texts in SNS. *Journal of Digital Contents Society, 19(1),* 51-58.
DOI : 10.9728/dcs.2018.19.1.51

[21] H. J. Lee & J. W. Kim. (2017). A Study on the Natural Language Processing(NLP) Technical and Standardization Trend. *Proceeding of Korea Institute of Communication Sciences,* 876-877.

이 후 영(Hoo-Young Lee) [정회원]

· 2002년 2월 : 우송대학교 컴퓨터과학과(공학사)
· 2017년 2월 : 공주대학교 대학원 멀티미디어공학과(공학석사)
· 2017년 3월 ~ 현재 : 공주대학교 대학원 컴퓨터공학과 박사과정
· 관심분야 : 인공지능, 빅데이터, 하둡, 정보보안
· E-Mail : hooyoung.paul.lee@gmail.com

박 구 락(Koo-Rack Park) [정회원]

· 1986년 2월 : 중앙대학교 전기공학과(공학사)
· 1988년 2월 : 숭실대학교 전자계산학과(공학석사)
· 2000년 2월 : 경기대학교 전자계산학과(이학박사)
· 1991년 4월 ~ 현재 : 공주대학교 컴퓨터공학부 교수
· 관심분야 : 인공지능, 경영정보, 영상처리, 전자상거래
· E-Mail : ecgrpark@kongju.ac.kr

김 동 현(Dong-Hyun Kim) [정회원]

· 1986년 2월 : 중앙대학교 전기공학과(공학사)
· 2005년 2월 : 공주대학교 컴퓨터멀티미디어공학과(공학석사)
· 2010년 2월 : 공주대학교 컴퓨터공학과(공학박사)
· 2016년 6월 ~ 현재 : ㈜정보소프트 기술이사
· 관심분야 : 인공지능, 영상처리, 지식관리, 시뮬레이션
· E-Mail : dhkim977@naver.com