

비지도학습 기반의 행정부서별 신문기사 자동분류 연구

김현종¹, 유승의^{1*}, 이철호², 남광우³

¹동아대학교 스마트거버넌스연구센터, ²한국과학기술원 기술경영학부, ³경성대학교 도시공학과

A Study on Automatic Classification of Newspaper Articles Based on Unsupervised Learning by Departments

Hyun-Jong Kim¹, Seung-Eui Ryu^{1*}, Chul-Ho Lee², Kwang Woo Nam³

¹Smart Governance Research Center, Dong-A University

²Business and Technology Management, Korea Advanced Institute of Science and Technology

³Department of Urban Planning & Engineering, Kyung Sung University

요약 행정기관은 정책 대응성을 제고하기 위해 빅데이터 분석에 관심을 기울이고 있다. 빅데이터 중 뉴스 기사는 정책 이슈와 정책에 대한 여론을 파악하는데 중요한 자료로 활용될 수 있다. 한편으로 새로운 온라인 매체의 등장으로 뉴스 기사의 생산은 급격히 증가하고 있어 문서 자동분류를 통해 기사를 수집할 필요가 있다. 그러나 기존 뉴스 기사의 범주와 키워드 검색방법으로는 특정 행정기관 및 부서별로 업무에 관련된 기사를 자동적으로 수집하는 것에 한계가 있었다. 또한 기존의 지도학습 기반의 분류 기법은 다량의 학습 데이터가 필요한 단점을 가지고 있다. 이에 본 연구에서는 행정부서의 업무특징을 포함한 분류사전을 활용하여 기사의 분류를 효과적으로 처리하기 위한 방법을 제안한다. 이를 위해 행정기관의 업무와 신문 기사를 Word2Vec와 토픽모델링 기법으로 부서별 특징을 추출하여 분류사전을 생성하고, 행정 부서별로 신문 기사를 자동분류 한 결과 71%정도의 정확도를 얻었다. 본 연구는 행정부서별 신문 기사를 자동분류하기 위해 부서별 업무 특징 추출 방법과 비지도학습 기반의 자동분류 방법을 제시하였다는 학문적·실무적 기여점이 있다.

Abstract Administrative agencies today are paying keen attention to big data analysis to improve their policy responsiveness. Of all the big data, news articles can be used to understand public opinion regarding policy and policy issues. The amount of news output has increased rapidly because of the emergence of new online media outlets, which calls for the use of automated bots or automatic document classification tools. There are, however, limits to the automatic collection of news articles related to specific agencies or departments based on the existing news article categories and keyword search queries. Thus, this paper proposes a method to process articles using classification glossaries that take into account each agency's different work features. To this end, classification glossaries were developed by extracting the work features of different departments using Word2Vec and topic modeling techniques from news articles related to different agencies. As a result, the automatic classification of newspaper articles for each department yielded approximately 71% accuracy. This study is meaningful in making academic and practical contributions because it presents a method of extracting the work features for each department, and it is an unsupervised learning-based automatic classification method for automatically classifying news articles relevant to each agency.

Keywords : Automatic Classification, Unsupervised Learning, Classification Glossary, Newspaper, Departments

이 논문은 2018년 대한민국 교육부와 한국연구재단의 지원을 받아 수행된 연구임 (NRF-2018S1A3A2075240)

*Corresponding Author : Hyun-Jong Kim(Dong-A Univ.)

email: juim0928@dau.ac.kr

Received May 29, 2020

Accepted September 4, 2020

Revised September 2, 2020

Published September 30, 2020

1. 서론

행정기관에서 빅데이터를 활용하여 정책 대응성을 높이기 위한 연구가 다양하게 이루어지고 있다. 이때 사용되는 데이터는 통계자료와 같이 수치나 일정한 형태를 갖춘 정형 데이터와 텍스트, 이미지, 영상 등과 같이 일정한 형식이 없는 비정형 데이터로 나눌 수 있다. 이중 비정형 데이터는 전체 데이터의 80%를 차지하고 있으며 온라인을 통해 증가 폭이 급격히 커지고 있지만 정형 데이터에 비해 아직 연구가 부족한 실정이다. 이러한 비정형 데이터 중 뉴스 기사, SNS, 각종 커뮤니티 사이트에서 발생하는 텍스트 데이터는 사람들의 말과 생각을 엿볼 수 있는 자료이다. 텍스트 분석을 통해 국민과 고객의 요구사항을 파악할 수 있기 때문에 정부 및 기업에서 다양한 방법으로 분석에 활용하고 있다. 텍스트 데이터를 분석하는 연구는 각종 온라인 서비스에서 발생하는 텍스트를 활용하고 있으며 이중에서 뉴스 기사는 여론을 파악하기 좋은 데이터이기 때문에 분석의 활용도가 높은 편이다. 뉴스 기사는 정책을 개발하기 위해 정책 이슈를 파악하는 방법 중의 하나로 행정기관에서 활용하고 있다 [1].

행정기관의 뉴스 기사를 수집하기 위한 방법은 기사를 범주별로 수집하거나 키워드 검색을 주로 사용하고 있다. 그러나 현재 언론 기사는 일반적으로 정치, 경제, 사회 등의 범주로 분류되고 있기 때문에 행정기관의 부서별 업무와는 일치하지 않는다. 예를 들어 일반적으로 신문기사는 정치, 경제 등으로 구분되어 있어 교통, 도시계획, 복지 등 기관 및 부서의 업무와는 분류 기준이 상이하여 기관과 부서의 업무에 맞게 분류하기 어렵다. 또한 키워드 검색을 통한 기사 수집은 특정 이슈를 파악하는데 적합하지만 기관 전체 또는 부서별 업무 전반에 해당하는 기사를 수집하기에는 한계가 있다. 이에 따라 매일 대량으로 발생하는 뉴스 기사들에서 부서 업무와 연관된 기사를 수집하는 것은 매우 어려운 실정이다.

이에 본 연구에서는 행정기관의 정책 대응성 제고를 위한 빅데이터 시스템에서 뉴스 기사를 해당 행정기관 및 부서별로 자동으로 분류하는 방법을 제안하고자 한다. 이 방법은 정책과정에 영향을 미치는 뉴스 기사의 수집과 분석을 위해서는 행정기관 및 부서의 업무에 적합한 기사를 분류하기 위해 학습 데이터가 필요 없는 방법이다. 연구의 착안점은 다음과 같다. 기사를 행정부서별로 분류하는 것은 학습 데이터 없이 범주화 시키는 것이므로 기존의 문서 자동분류 방법으로는 해결할 수 없다. 이

러한 문제 해결을 위한 방법으로 일종의 사전(dictionary)과 같이 미리 정의된 분류 기준을 이용할 수 있을 것이다. 즉, 행정부서의 고유 업무로부터 분류 기준이 되는 특징을 추출하여 분류사전을 생성하고 이 분류사전을 이용하여 학습 데이터 없이 신문 기사를 부서별로 자동분류하는 것이다. 본 연구에서 제안하는 방법은 B시의 행정부서별로 신문 기사를 자동분류하기 위한 과정을 통해 구체적인 방법과 절차를 설명하였다. 사전 생성을 위해 텍스트 마이닝 기법을 이용하여 부서별 업무 특징을 추출하였고 생성된 사전을 활용하여 학습 데이터가 필요 없는 비지도 학습기반으로 신문 기사 자동분류를 실시하였다.

2. 이론적 배경

문서 자동분류는 컴퓨터를 이용하여 유사한 문서를 같은 집단으로 나누는 기법이다[2]. 문서 자동분류는 일반적으로 분류(classification) 기법을 이용하는 방법과 군집화(clustering)를 이용하는 방법으로 나누어진다. 분류는 사전에 정의된 범주에 따라 문서를 배치하는 방법이고 군집화는 사전 정의된 범주 없이 유사한 문서를 집단화하는 방법이다[3].

일반적으로 문서 자동분류를 위해 주로 지도학습 방법이 사용되었다. 지도학습 기반의 문서 자동분류는 레이블이 있는 문서를 전처리와 특징추출을 거쳐 지도학습 알고리즘으로 학습을 시킨 후 미분류된 문서를 분류하는 과정을 거치게 된다. 지도학습 기반의 연구는 결정트리, KNN(K-Nearest Neighbor classification) 기법, 나이브 베이즈 (Naive Bayes) 기법, SVM(Support Vector Machine) 방법, 딥러닝(deep learning) 기법 등이 사용되고 있다. 최근 사용되는 인공신경망과 딥러닝 기법은 보다 정확한 결과를 보여주지만 더욱 많은 학습 데이터를 요구하기 때문에 이러한 기법을 사용하기 위해 학습 데이터의 확보가 더욱 중요해지고 있다. 지도학습은 데이터에 레이블을 작성하기 위해 많은 시간과 비용이 소요되기 때문에 이를 해결할 수 있는 비지도학습 기반의 자동분류에 대한 연구가 더욱 요구된다.

비지도학습 방식은 레이블이 없는 데이터를 학습하여 패턴, 특징 등의 규칙성을 찾아 유의미한 정보를 얻고자 할 때 이용되는 기계학습 방법이다. 비지도학습을 통한 문서분류는 문서들 사이의 유사성을 측정하여 같은 특징을 가지는 그룹으로 군집화하게 된다. 군집화는 비슷한

문서를 유사한 주제로 구분하기에 효과적인 방법이지만 정해진 범주로 분류하기 어렵다.

이에 따라 레이블이 없는 문서의 범주화를 위한 연구들이 시도되고 있다. 새로운 방법으로는 기존의 비지도학습에 준지도학습의 방법을 가미하거나 분류 기준 또는 규칙을 제시하는 방법을 사용하고 있다. 본 연구에서도 분류의 기준이 되는 분류사전을 활용하여 레이블이 없는 기사를 비지도학습으로 범주화 할 수 있는 방법을 제안한다.

3. 연구방법 및 실험

3.1 연구방법 개요

본 연구에서 제안하는 행정기관 및 부서별로 적합한 기사의 자동분류는 Fig. 1 같은 절차 및 방법으로 구성되었다. 연구방법은 크게 분류의 기준이 되는 사전 생성과 사전을 이용하여 기사를 부서별로 자동분류하는 과정으로 이루어져 있다.

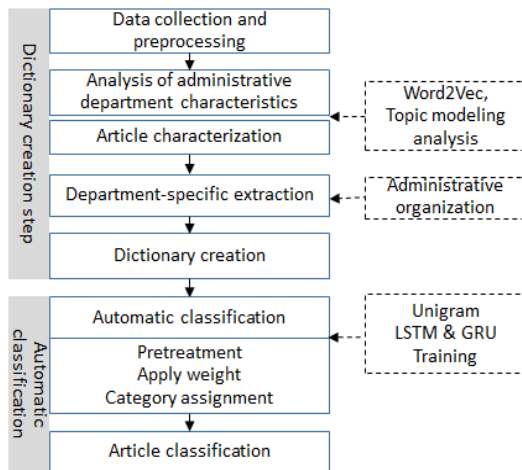


Fig. 1. Research procedures and methods

본 연구에서 제안하는 연구방법은 실제 지방정부의 행정기관을 대상으로 기사 자동분류를 실험하기 위해 B시의 행정부서를 대상으로 신문 기사를 분류하였다. B시는 광역지자체로 시에는 수많은 행정부서 및 산하기관을 가지고 있지만 본 연구에서는 시청 내 실국본부급 부서로 한정하였으며, 신문기사 1년 치를 분류하였다.

3.2 행정부서 특징 분석

행정부서별 특징을 파악하기 위해 B시 홈페이지에 게재된 실국본부에 소속된 직원의 업무를 수집하여 부서별 업무 데이터를 확보하였다. 부서 업무의 특징을 추출하기 위해 Word2Vec 기법을 활용하였다. Word2Vec는 인공신경망을 활용하여 대량의 텍스트 데이터를 단어의 의미와 구문의 정보를 벡터로 효율적으로 표현할 수 있는 방법이다[4]. 이렇게 표현된 단어의 벡터는 비슷한 의미를 가진 단어를 찾아내는데 활용할 수 있다. 이에 본 연구에서도 Word2Vec를 이용해 부서별 업무에 등장하는 단어들을 벡터로 얻은 결과를 통해 유사도를 분석하여 대표 업무의 단어를 추출하였다.

3.3 신문기사 특징성 분석

신문기사 특징 분석은 업무분장의 분석만으로 부족한 부서별 특징을 보완하기 위한 방법이다. 업무분장에는 구체적으로 나타나지 않지만 신문에서 다루고 있는 내용을 파악하기 위해 업무와 연관 있는 기사에 사용되는 업무 특징을 파악하는 것이다. 신문기사의 특징을 분석하기 위해 중앙지 1개와 해당 지역신문 1개의 정치, 경제, 사회 분야의 기사 6개월 치 42,000여 건을 수집한 후, 텍스트 마이닝을 이용하여 특징을 추출하였다. 수집된 기사는 형태소 분석을 통해 명사를 추출하였고 대명사, 숫자, 1글자 단어 등 불용어를 제거하였다. 핵심주제어를 도출하기 위해 빈도 분석을 실시하여 상위 100개의 단어를 추출하였다. 또한 기사의 주요 주제를 추출하기 위해 LDA(Latent Dirichlet Allocation) 토픽모델링을 활용하여 기사를 주제별로 군집화 하였다. 토픽모델링은 이슈 토픽을 분석하기 위해 문서 집합의 자주 등장하는 단어로부터 잠재되어 있는 주제를 발견하는 방법이며 비지도 학습의 대표적 군집화 기법이기도 하다. 주제 토픽수는 50개에서 200개로 늘려가며 토픽의 결과를 분석한 결과 200개의 토픽수에서 다양한 주제로 분류되는 것을 확인할 수 있었으며 이때 토픽별로 주요 단어를 주제별로 파악하여 기사의 특징으로 추출하는데 활용하였다. 이를 위해 사용된 프로그램은 통계 프로그램 R이며, KoNLP, tm, lda, topicmodels 패키지를 사용하였다. 핵심 주제는 사전에 포함될 단어의 기준이 되며 토픽모델링에서 나온 주제와 포함된 키워드를 행정부서와 비교하게 된다. Table 1은 토픽 모델링 결과를 비슷한 주제별로 정리한 것이다.

Table 1. Result of topic modeling analysis

Similar topics	Number of topics	Similar topics	Number of topics
President / Blue House	5	real estate	3
National Assembly / Political Party	10	International	5
Prosecution / Police	10	Incident / accident	11
education	6	Agency	6
Labor / union	7	Court / lawsuit	4
Trump / United States	3	Medical treatment	6
North Korea	7	Employment	2
Enterprise	5	female	2
finance	6	Samsung	2
broadcast	4	industry	8
Environment / Weather	5	Politics other	7
area	8	Economy other	13
Concentrated fisheries	3	Social other	36
Product / PR	8		

대표 주제는 토픽 중 대통령/청와대, 국회/정당, 북미 회담과 관련 주제는 정치 및 외교 관련 사항으로 개별적인 행정부서와는 직접적인 연관성이 부족하여 특징 분석에서 제외하였다. 나머지 주제를 행정부서와 연관된 주제로 정리한 결과는 Table 2와 같다. 토픽 키워드와 업무의 연관성 파악을 위해 정책 관련 전문가 3인의 검토를 거쳐 분류하였다.

Table 2. Relevance of Topic keywords and administrative departments

Topic Keywords	Administrative Agencies
Civic groups, polls	Citizen Happiness & Communications Headquarters
FTC, regulations, union election, disciplinary action	Audit Committee
Local Yeta, Oil Prices, Religious Taxation	Planning & Coordination Office
Product safety, foot and mouth disease, illegal detection, nuclear power plant safety	Public Safety Office
Business construction, housing business, safety construction	Urban Planning Office
Real estate prices, apartment sales	Balanced Urban Regeneration Bureau

Female art, gourmet, travel, coffee	Culture & Sports Bureau
Hospital, pension, treatment, treatment, epidemic, mental illness, tobacco	Welfare & Health Bureau
Female hate, family, comfort women, abortion, sexual assault coach	Women & Family Affairs Bureau
Universities, education meals/accidents/support/schools, worker safety/wages, union minimum wages/strikes, ward offices, residents, events, kindergartens, government countermeasures/ policies, regions, obituaries, population, human rights	Administrative Management Bureau
Bus strike, Asiana sale, taxi couple, transportation, drunk driving, vehicle	Transportation Headquarters
Marine ships, ports	Logistics Bureau
Economic Outlook, Russia, US-China Trade, Vietnam	Growth Strategy Headquarters
Employment youth, employment statistics, income support, small business market	Employment & Economy Office
Finance, Investment, Corporate Sales/Generation/Support/ Permit, Industrial Support, Energy Power, Bank Interest/Loan/Service, Electric Vehicle, Stock, Product Launch	Future Industry Bureau
Weather, fine dust air / reduction, environmental plastic, national park	Environment Policy Office
Environment Nakdong River	Water Policy Bureau
Marine fishery, agri-food	Maritime Affairs, Fisheries Bureau
Wildfire, sewol, fire	Fire Marshal

3.4 분류사전 생성

행정기관과 뉴스 기사의 특성을 행정기관의 조직체계에 맞춰 분류 기준으로 사용할 키워드를 리스트로 정리하여 분류사전을 생성하였다. 행정부서 중 실국본부의 다른 부서와 구분되는 고유의 업무특성을 나타낼 수 있는 키워드로 정리한 결과는 Table 3과 같다.

3.5 신문기사 자동 분류

앞서 생성한 분류사전을 활용하여 뉴스 기사를 분류하고 정확도를 검증하였다. 검증을 위한 신문기사는 특성 분석에 사용되지 않은 1개 중앙지 1년간의 기사 41,275 개를 사용하였다. 분류는 비지도학습을 기반으로 행정부서와 기사의 특성을 반영하여 도출된 사전의 키워드에 가중치를 적용하여 기사를 분류하였다. 자동분류기는 파이썬을 이용하였으며 기사를 전처리하여 TF-IDF와 Word2Vec를 이용하여 가중치를 계산하여 기사를 행정부서별 분류하였다. 앞서 제외된 정치, 북한 관련 기사는 해당 부서가 없는 것으로 처리되도록 하였다.

Table 3. Administrative Agencies work characteristics

Administrative Agencies	work characteristics
Citizen Happiness & Communications Headquarters	Media, public relations, conflict, press releases, correction, analysis
Audit Committee	Audit, contract, budget, integrity, corruption
Planning & Coordination Office	Correction, Policy, Planning, Personnel, Recruitment, Legal, Information, Big Data, Advancement, BSC, Budget, Finance, Expense, Taxation, Organization, Expenditure
Public Safety Office	Safety, disaster, nuclear power plant, maintenance, inspection, surveillance, investigation
Urban Planning Office	City planning, establishment, urban development, road, construction, land, real estate
Balanced Urban Regeneration Bureau	Urban Regeneration, Redevelopment, New Deal, Housing, Architecture
Culture & Sports Bureau	Culture, art, physical education, cultural properties, tourism, festivals, content, sports, leisure
Welfare & Health Bureau	Welfare, medical care, hospital, elderly, epidemic, disabled, tobacco
Women & Family Affairs Bureau	Women, family, multicultural, childbirth, children, youth, sexual violence
Administrative Management Bureau	Local difference, human rights, labor, local tax, education, school, staff, civil complaint
Transportation Headquarters	Transportation, bus, taxi, subway, public transportation
Logistics Bureau	Logistics, ports, railways, cargo, infrastructure
Growth Strategy Headquarters	Economy, science and technology, diplomacy, youth, employment, industry-university cooperation
Employment & Economy Office	Jobs, entrepreneurship, R & D, investment, market, village company, investment, industrial complex
Future Industry Bureau	Industry, manufacturing, automobile, shipbuilding, textile, finance, smart city, energy, growth, trade
Environment Policy Office	Environment, weather, fine dust, waterworks, garbage, parks, green areas
Water Policy Bureau	Water, rivers, Nakdong River, sewage, discharge, boom
Maritime Affairs, Fisheries Bureau	Marine, cruise, fisheries, agriculture, livestock, fishing
New Airport Promotion Headquarters	New Airport, Gimhae, Gadeokdo
Fire Marshal	Fire, disaster, forest fire, rescue, firefighting, prevention

분류사전을 이용하여 분류된 결과의 정확도를 검증하기 위해서는 새로운 방법이 필요하다. 본 연구의 자동분류 실험에 사용된 신문기사에는 레이블이 없기 때문에 기사가 해당 부서로 정확히 분류되었는지 확인하기 어려우며 전체를 수작업으로 검토하는 것도 무리가 있다. 이에 따라 분류의 정확도를 간접적으로 측정하기 위해 비지도학습 기반의 대표적인 군집화 방법인 LDA(Latent Dirichlet Allocation) 결과와 비교하였다. LDA는 문서를 유사한 주제별로 그룹화하는 성능이 우수한 방법이므로 자동분류로 분류된 기사가 부서별로 유사한 주제로 모였는지 비교할 수 있다. 이를 이용하여 자동분류의 정확도를 측정하였다. 신문기사 분류의 실험 결과는 Fig. 2와 같이 약 71%의 정확도를 보였다. 실험 결과에서 나타난 정확도는 분류된 기사가 얼마만큼 정확하게 부서별로 할당이 되었는지는 보여주지는 못하지만 자동분류된 기사가 동일한 주제로 잘 모였는지는 알 수 있다. 부서별로 모인 주제가 비슷한 주제로 모인 정도가 71%를 나타내는 것을 뜻한다. 이는 부서별 업무를 기준으로 기사의 동

일한 주제를 가질 가능성이 높기 때문에 이와 같은 방법을 활용하여 간접적인 정확도를 계산하였다. 자동분류 연구의 정확도는 얼마 이상이면 합격이라는 기준이 없다. 기존 기술보다 향상된 성능을 가지도록 개발하는 것이 목적이기 때문에 본 연구에서 새롭게 제안한 방법이 향후 연구의 가능성을 제시하였다는 것에 의미가 있다.

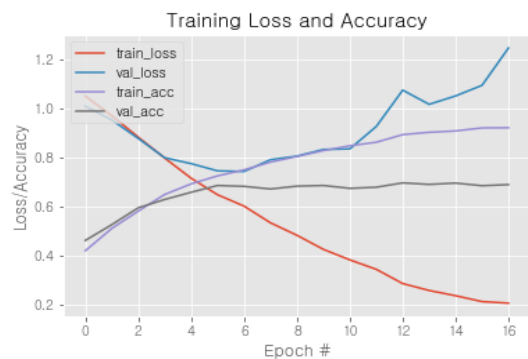


Fig. 2. Classification result

4. 결론

기존 뉴스 기사는 언론사가 정한 범주로 분류되어 있어 행정기관과 부서 업무와는 맞지 않아 정책 및 업무별로 별도로 수집해야하는 문제가 있었다. 또한 기존 신문기사의 분류방법도 지도학습 기반의 분류 기법을 사용하고 있어서 대량의 학습 데이터가 필요하였다. 이에 본 연구에서는 분류의 기준이 되는 분류사전을 이용한 비지도 학습 기반의 문서 자동분류 방법을 제안하였다. 이를 위해서는 학습 데이터 없이 기사를 분류하기 위한 분류기법이 필요하였으며 문서의 자동분류를 위해 행정부서의 특성이 포함된 사전을 이용하였다. 행정부서의 특성이 포함된 사전은 부서별 업무를 Word2Vec로 분석해 특성을 추출하였고 신문기사를 토픽모델링으로 분석해 나온 토픽 200개로부터 부서와 연관된 특성을 추출한 결과를 이용하여 생성하였다. 이렇게 도출한 부서별 특성을 사전으로 활용하여 신문기사를 행정부서별로 분류한 결과는 학습시 90% 이상의 정확도를 보였으며, 실제 분류에서는 71%의 정확도를 나타내었다.

본 연구에서는 행정기관의 업무 특성을 포함하는 사전을 활용한 뉴스 기사 자동분류 방법을 제안하였다. 이는 행정기관에서 정책 대응성을 제고하기 위한 빅데이터 분석을 활용하기에 앞서 데이터의 수집과 관리가 필요하다. 특히 본 연구에서 주목한 점은 학습 데이터가 없는 비정형 데이터인 텍스트 및 문서를 해당 부서별로 분류할 수 있는 방법을 제시하였다. 이를 위해 분류사전의 생성과 활용에 대한 새로운 방법을 제시하였다는 학문적 기여점이 있다. 또한 행정기관이 신문기사를 수집하여 정책이슈를 파악하고 정책과정에 반영하는데 활용이 가능한 실무적인 방법을 제시하였다.

이번 연구에서 실험한 자동분류기는 하나의 신문 기사를 하나의 부서로 할당하게 하였다. 실제로 어떠한 기사는 2개 이상의 부서와 연관된 기사가 있을 수 있으며 해당 부서 모두에게 전달해야 하는 경우가 있을 수 있다. 향후 연구에서는 일정 수준 이상의 연관성을 가지는 기사에 해당 부서들에게 동시에 할당할 수 있는 기준을 마련하고 기술적 보완이 필요한 것으로 보인다.

References

[1] T. Kurtoglu, I. Y. Tumer, D. C. Jensen, "A functional failure reasoning methodology for evaluation of

conceptual system architecture", *Research in Engineering Design*, Vol.21, No.4, pp.209-234, Oct. 2010.

DOI: <http://doi.org/10.1007/s00163-010-0086-1>

[2] S. G. Hong, H. J. Kim, N. R. Kim, "Development of a Co-creation Model for the Social Problem Resolution", *Information*, Vol.19, NO.8, pp.3395-3400, Aug. 2016.

[3] J. S. Kim, Y. J. Kim, H. J. Mun, Y. T. Woo, "A Feature Selection Technique for an Efficient Document Automatic Classification", *Journal of Information Technology Applications & Management*, Vol.8, No.1, pp.117-128, Jul. 2001.

[4] D. Koller, M. Sahami, "Hierarchically Classifying Documents Using Very Few Words", *Proceedings of the Fourteenth International Conference on Machine Learning*, ICML 97, pp.170-178, Jul. 1997.

[5] J. Y. Jeong, K. H. Mo, S. W. Seo, C. Y. Kim, H. D. Kim, P. S. Kang, "Unsupervised Document Multi-Category Weight Extraction based on Word Embedding and Word Network Analysis : A Case Study on Mobile Phone Reviews", *Journal of the Korean Institute of Industrial Engineers*, Vol.44, No.6, pp.442-451, Dec. 2018.

김 현 종(Hyun-Jong Kim)

[정회원]



- 2012 2월 : 동아대학교 경영정보학과 (경영학석사)
- 2020년 2월 : 동아대학교 경영정보학과 (경영학박사)
- 2020년 3월 ~ 현재 : 동아대학교 스마트거버넌스연구센터 전임연구원

<관심분야>

경영정보, 텍스트 마이닝, 스마트 거버넌스

유 승 의(Seung-Eui Ryu)

[정회원]



- 2012 2월 : 동아대학교 경영정보학과 (경영학석사)
- 2014년 5월 : University of Texas at El Paso 박사 (Information Systems)
- 2018년 9월 ~ 현재 : 동아대학교 스마트거버넌스연구센터 전임연구원

<관심분야>

국제경영학, 빅데이터, 인공지능, 텍스트마이닝

이 철 호(Lee, Chul Ho)

[정회원]



- 2002년 2월 : Pusan National University (MA in Business Administration)
- 2012년 8월 : the University of Texas at Dallas (Ph.D. in Management Science)
- 2013년 10월 ~ 2016년 1월 : 하얼빈공업대학교 부교수
- 2016년 2월 ~ 현재 : 카이스트 기술경영학부 조교수

<관심분야>

Economics of Information Systems, Platform Economics

남 광 우(Kwang-Woo Nam)

[정회원]



- 1998년 2월 : 부산대학교 지형정보협동과정 (공학석사)
- 2001년 8월 : 부산대학교 지형정보협동과정 (공학박사)
- 2003년 3월 ~ 현재 : 경성대학교 도시공학과 교수

<관심분야>

스마트도시계획, 도시공간모델링, 리빙랩