



베타 회귀분석과 R 텍스트 마이닝을 이용한 특허 마이닝

Patent Mining using Beta Regression and R Text Mining

저자 (Authors)	전성해 Sunghae Jun
출처 (Source)	한국지능시스템학회 논문지 28(4) , 2018.8, 383-387 (5 pages) Journal of Korean Institute of Intelligent Systems 28(4) , 2018.8, 383-387 (5 pages)
발행처 (Publisher)	한국지능시스템학회 Korean Institute of Intelligent Systems
URL	http://www.dbpia.co.kr/Article/NODE07520679
APA Style	전성해 (2018). 베타 회귀분석과 R 텍스트 마이닝을 이용한 특허 마이닝. 한국지능시스템학회 논문지, 28(4), 383-387.
이용정보 (Accessed)	성균관대학교 자연과학캠퍼스 115.***.238.89 2018/09/17 16:40 (KST)

저작권 안내

DBpia에서 제공되는 모든 저작물의 저작권은 원저작자에게 있으며, 누리미디어는 각 저작물의 내용을 보증하거나 책임을 지지 않습니다. 그리고 DBpia에서 제공되는 저작물은 DBpia와 구독 계약을 체결한 기관소속 이용자 혹은 해당 저작물의 개별 구매자가 비영리적으로만 이용할 수 있습니다. 그러므로 이에 위반하여 DBpia에서 제공되는 저작물을 복제, 전송 등의 방법으로 무단 이용하는 경우 관련 법령에 따라 민, 형사상의 책임을 질 수 있습니다.

Copyright Information

Copyright of all literary works provided by DBpia belongs to the copyright holder(s) and Nurimedia does not guarantee contents of the literary work or assume responsibility for the same. In addition, the literary works provided by DBpia may only be used by the users affiliated to the institutions which executed a subscription agreement with DBpia or the individual purchasers of the literary work(s) for non-commercial purposes. Therefore, any person who illegally uses the literary works provided by DBpia by means of reproduction or transmission shall assume civil and criminal responsibility according to applicable laws and regulations.



베타 회귀분석과 R 텍스트 마이닝을 이용한 특허 마이닝

Patent Mining using Beta Regression and R Text Mining

전성해

Sunghae Jun

청주대학교 빅데이터통계학과

Department of Big Data and Statistics, Cheongju University

요약

개발된 기술에 대한 특허는 숫자, 문자, 그림 등으로 이루어진 문서형식이다. 특허 마이닝은 대규모 특허문서 데이터로부터 기술과 관련된 다양한 지식을 추출하는 도구와 방법이다. 문서는 대부분 텍스트로 구성되어 있기 때문에 특허 마이닝에서는 텍스트 데이터를 처리하고 분석할 수 있는 텍스트 마이닝이 필요하다. 텍스트 마이닝을 지원하는 프로그래밍 언어로 본 논문에서는 R을 사용한다. R은 텍스트 마이닝 뿐만 아니라 대부분의 통계분석과 기계학습 알고리즘을 지원한다. 제안 방법에서 사용되는 베타 회귀분석도 R의 통계 패키지를 이용하여 수행된다. 반응변수가 0에서 1사이의 값을 갖는 베타 회귀분석의 특성을 이용하여 본 연구에서는 특허문서로부터 추출된 키워드 사이의 기술 연관성을 찾는 방법을 제안한다. 제안 모형의 성능평가를 위하여 실제 특허문서를 이용한 실험을 수행한다.

키워드 : 특허마이닝, 기술예측, 베타회귀분석, R 데이터언어, 텍스트마이닝

Abstract

Patents for developed technologies are in the form of documents consisting of numbers, texts and pictures. Patent mining means tools and methods for extracting various knowledge related to technology from large-scale patent document data. Since documents are mostly text, patent mining requires text mining to process and analyze text data. Also, R is used as a programming language that supports text mining. R supports almost all statistical analysis and machine learning algorithms as well as text mining. The beta regression analysis used in the proposed method is also performed using the R statistical package. In this paper, we propose a method to find the technological relation between patent keywords extracted from patent documents by using the characteristics of beta regression analysis with response variables between 0 and 1. Experiments using real patent documents are performed to evaluate the performance of the proposed model.

Key Words : Patent Mining, Technology Forecasting, Beta Regression, R Data Language, Text Mining

Received: Jul. 31, 2018

Revised: Aug. 7, 2018

Accepted: Aug. 8, 2018

[†]Corresponding authors
shjun@cju.ac.kr

1. 서론

기업은 자사가 개발한 기술의 권리를 보호받기 위하여 우리나라를 비롯한 전세계의 특허청을 대상으로 해당 기술에 대한 특허를 출원하고, 심사과정을 통하여 최종적으로 등록특허로 만든다. 4차 산업혁명으로 인하여 기술에 대한 기업의 의존도가 더욱 커지고 있기 때문에 자사의 기술 경쟁력을 확보하기 위하여 연구개발 분야에서는 기술에 대한 분석에 많은 관심을 보인다. 현재 목표 기술 분야에 대한 기술 분석은 크게 2가지 접근방법으로 나뉜다 [1]. 첫째는 해당 기술 분야의 전문가 집단에 의해 기술 분석을 수행하는 델파이조사(Delphi survey)이다 [2]. 즉 해당 기술 분야에 대한 분석을 위한 설문문항을 작성하고 이를 전문가 집단에게 반복적으로 제공하고 회신하여 모든 전문가들의 의견이 수렴되는 지점에서 분석결과를 도출하는 방법이다. 델파이조사는 해당분야 전문가의 지식을 반영할 수 있지만 이것이 주관적이라는 문제점이 있다. 이에 비하여 정량적 기술 분석은 객관적인 특허 데이터를 분석함으로써 전문가의 주관적 지식이 반영되지 않은 데이터에 기반 한 기술 분석이 이루어진다. 정량적 특허분석은 객관적 기술 분석이 가능하다는 장점으로 최근 많은 연구가 이루어지고 있다 [3-5]. 즉 다양한 빅데이터 분석기법을 사용하여 해당 기술 분야의 특허문서를 수집하고 분석하여 필요한 기술 분석을 수행한다 [6]. 기술은 시간에 따라 진화하기

This is an Open-Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

때문에 기술의 성장을 반영한 통계적 모형화가 필요하다. 기존의 특허분석은 기술의 성장을 반영하지 않은 정량적 분석이었기 때문에 기술 분석 결과의 해석에 한계가 있었다. 본 연구에서는 이와 같은 기술의 성장과정을 반영한 통계적 모형을 사용한 정량적 기술 분석 방법으로서 특허 마이닝(patent mining)을 제안한다. 또한 제안방법의 성능평가를 위하여 실제 특허데이터를 수집하고 분석하였다. 분석결과를 통하여 제안방법의 실제적용 가능성을 확인하였다. 2장에서는 제안하는 방법에 대한 설명을 나타내고 3 장에서는 실험 및 결과를 그리고 마지막 장에서는 결론과 향후 연구과제에 대하여 알아본다.

2. 베타회귀분석과 R 텍스트 전처리를 이용한 특허 마이닝

본 연구에서는 베타회귀분석(beta regression analysis)과 R 텍스트 전처리를 이용한 특허 마이닝 방법을 제안한다. 특허 마이닝을 위하여 우선 목표기술 분야에 해당하는 특허 문서를 수집해야 한다. 특허문서의 수집을 위하여 키워드 검색식 등이 사용된다 [7]. 베타회귀모형과 같은 통계적 분석을 위하여 수집된 특허문서는 전처리 과정을 거친다. 본 연구에서는 특허 데이터의 전처리를 위하여 다양한 텍스트 마이닝 방법을 사용한다 [8]. 실제 특허문서의 전처리를 위한 텍스트 마이닝을 위하여 R 데이터 언어와 R이 제공하는 텍스트 마이닝 패키지를 사용한다 [9-11]. 본 연구에서는 수집된 최종 특허문서의 R 텍스트 전처리 과정을 통하여 특허문서에 포함되었던 특허 키워드를 추출한다. 추출된 키워드에 대한 각 특허 문서에 나타난 빈도를 베타회귀분석 모형의 설명변수(explanatory variable) 및 반응변수(response variable)로 사용한다. 이 과정에서 특허 마이닝을 위한 입력변수로 사용되는 기술특허 키워드는 기술의 추세에 따라 변화하는 경향을 나타낸다. 즉 기술 키워드는 아래와 같이 기술의 성장에 따라 성장곡선(growth curve)를 갖게

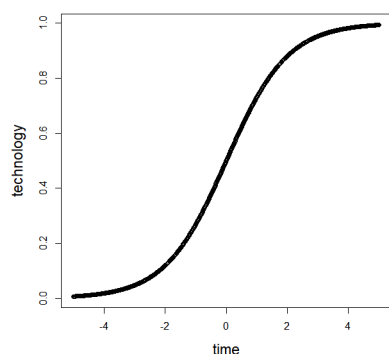


그림 1. 기술성장곡선

Fig. 1. Technology growth curve

된다 [1].

기술성장곡선(technology growth curve; TGC)은 0에서 1사의 값을 이용하여 해당 기술의 도입에서 성숙까지를 나타낸다. 본 연구에서는 목표기술에 해당하는 기술키워드 Y 를 다음과 같이 0과 1사이의 값으로 바꾼다.

$$-\infty < Y < \infty \rightarrow Y_{TGC} = \frac{1}{1 + \exp(-Y)} \quad (1)$$

Y_{TGC} 는 제안 방법의 베타회귀분석 모형의 반응변수로 사용된다. 일반적으로 베타회귀분석은 다음과 같이 정의되는 베타 분포(beta distribution)에 기반 한다.

$$Y \sim B(\alpha, \beta), \quad 0 < Y < 1, \quad \alpha > 0, \quad \beta > 0 \quad (2)$$

α 와 β 는 아래와 같이 정의되는 베타 확률변수 Y 의 확률분포를 결정하는 2개의 모수를 나타낸다 [12]. Y 는 0과 1사의 값을 가지며 이는 특허키워드의 성장을 나타낸다.

$$f(Y=y) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} y^{\alpha-1} (1-y)^{\beta-1} \quad (3)$$

여기서 $\Gamma(\cdot)$ 는 감마함수(gamma function)를 나타낸다. 또한 베타확률변수 Y 의 평균 $E(Y)$ 와 분산 $Var(Y)$ 는 다음과 같이 계산 된다 [12].

$$E(Y) = \frac{\alpha}{\alpha + \beta}, \quad Var(Y) = \frac{\alpha\beta}{(\alpha + \beta)^2(\alpha + \beta + 1)} \quad (4)$$

이와 같은 베타확률분포를 이용하여 회귀모형을 구축하기 위해서 다음과 같은 모수변환(parameter transformation)이 필요하다 [13].

$$A = \frac{\alpha}{\alpha + \beta}, \quad B = \alpha + \beta, \quad 0 < A < 1, \quad B > 0 \quad (5)$$

이 변환을 통하여 모수 A 와 B 를 갖는 베타확률분포는 다음과 같다.

$$f(Y=y) = \frac{\Gamma(B)}{\Gamma(AB)\Gamma((1-A)B)} y^{AB-1} (1-y)^{(1-A)B-1}, \quad 0 < y < 1 \quad (6)$$

이 때 확률변수 Y 의 기댓값과 분산은 다음과 같이 나타낸다.

$$E(Y) = A, \quad Var(Y) = \frac{A(1-A)}{1+B} \quad (7)$$

n 개의 특허키워드 빈도 데이터를 y_1, y_2, \dots, y_n 로 나타내면 y_i 는 다음과 같은 베타확률분포를 따른다[14].

$$y_i \sim B(A_i, B), \quad i = 1, 2, \dots, n \quad (8)$$

이를 통하여 본 연구의 베타회귀분석모형은 다음과 같다.

$$g(A_i) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k \quad (9)$$

여기서 x_1, x_2, \dots, x_k 는 y 를 설명하는 k 개의 특허 키워드를 나타낸다. 최종적으로 본 연구에서 제안하는 특허 마이닝 절차는 다음과 같다.

(Step1) 목표 기술에 해당하는 특허 문서 검색 및 수집

(1-1) 목표 기술을 설명하는 키워드 검색식 작성

(1-2) 키워드 검색식을 이용하여 특허 데이터베이스로부터 해당하는 특허 수집

(Step2) 특허 문서의 전처리

(2-1) 수집된 특허문서로부터 유효특허 선정

(2-2) R 데이터언어 및 텍스트 마이닝 패키지를 이용하여 특허 키워드 추출

(Step3) 베타회귀분석을 이용한 특허 키워드 분석

(3-1) 추출된 키워드를 이용하여 키워드 빈도 데이터 생성

(3-2) 키워드 빈도 데이터를 이용한 베타회귀 모형구축

(Step4) 추정된 모수를 이용한 기술 연관 계층도 구축

(4-1) 회귀계수를 이용한 연관성 파악

(4-2) 유의한 회귀계수 중심으로 기술 계층도 생성

최종적으로 구축된 기술 연관 계층도를 이용하여 해당 기술 분야의 연구개발 기획 등을 위한 기술경영 활동에 사용된다. 다음 절에서는 제안된 방법의 실제 적용을 위하여 애플의 기술특허를 이용한 특허 마이닝 실험을 수행한다.

3. 실험 및 결과

제안하는 기술예측 방법론의 실제 적용을 위하여 애플이 지금까지 출원, 등록한 전체특허를 수집하고 분석하였다. 2012년까지 출원, 등록한 수집특허 중에서 총 8,199건의 유효특허를 선정하였다. R 텍스트 전처리를 통하여 다음과 같이 총 17개의 특허 키워드를 추출하였다; audio, computer, content, data, device, digital, display, electronic, image, information, interface, memory, mobile, system, text, user, video. 애플의 대표 키워드는 이 분야 전문가의 도움을 받아 결정하였다. 먼저 추출된 애플의 기술 키워드 중에서

'Data'를 베타회귀모형의 반응변수로 사용하였다. 일반적으로 애플의 스마트 기기를 위한 기술의 가장 기본이 되는 키워드가 데이터이기 때문이다. 표 1은 'Data'를 반응변수로 하고 나머지 키워드를 설명변수로 설정한 베타회귀분석 결과이다.

표 1. 베타회귀분석 결과: 반응변수=data

Table 1. Beta regression result: response variable=data

Keyword	Estimate	S.E.	p-value
audio	0.0452	0.0100	0.0000
computer	0.0541	0.0109	0.0000
content	-0.0409	0.0134	0.0023
device	0.0268	0.0074	0.0003
digital	-0.0217	0.0146	0.1375
display	-0.0565	0.0086	0.0000
electronic	-0.0860	0.0101	0.0000
image	0.0498	0.0087	0.0000
information	0.0154	0.0113	0.1730
interface	-0.0315	0.0139	0.0240
memory	0.1190	0.0094	0.0000
mobile	0.0273	0.0179	0.1283
system	0.0447	0.0079	0.0000
text	-0.0638	0.0165	0.0001
user	-0.0021	0.0089	0.8131
video	0.0399	0.0104	0.0001

표 1에서 제공하는 분석결과를 통하여 'Data'에 통계적으로 유의한 영향을 미치는 특허 키워드는 audio, computer, content, device, display, electronic, image, interface, memory, system, text, video 임을 확인하였다. 이들 키워드 변수에 대한 유의확률이 0.05보다 작게 나왔기 때문이다. 이 분석의 결과에 대한 성능평가를 확인하기 위하여 잔차(residual), 쿡의 거리(Cook's distance), 예측값(predicted values)의 분포 등을 그림 2에 나타내고 있다.

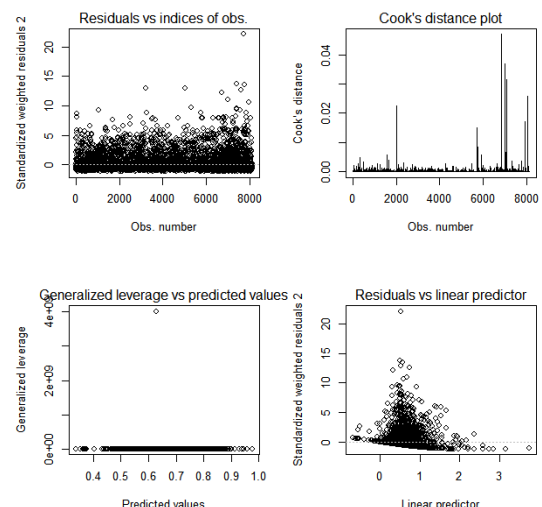


그림 2. 모형 평가 및 예측: 반응변수=data

Fig. 2. Model evaluation and prediction: response variable=data

그림 2의 각각의 결과를 통해 분석결과에 대한 타당성을 확인 할 수 있었다. 즉, 각 특허문서에 대한 잔차(residual)도 0을 중심으로 퍼져있고 Cook의 거리를 표시한 산점도 역시 제안모형의 성능에 대한 타당성을 확인하여 주고 있다. 본 연구에서는 애플의 'Data' 키워드에 영향을 미치는 설명변수들 중에서 'Audio', 'Video' 그리고 'Text'를 선택하여 각 키워드를 반응변수로 하는 베타회귀분석을 수행하였다. 즉 애플의 데이터 기술에 영향을 미치는 소리, 동영상, 그리고 문자 기술들에 대한 영향력을 확인하기 위함이다. 표 2는 'Audio'를 반응변수로 하고 나머지를 설명변수로 하는 베타회귀분석의 결과를 나타내고 있다.

표 2. 베타회귀분석 결과: 반응변수=audio

Table 2. Beta regression result: response variable=audio

Keyword	Estimate	S.E.	p-value
computer	-0.0100	0.0071	0.1626
content	-0.0122	0.0090	0.1761
device	0.0083	0.0049	0.0900
digital	0.1121	0.0103	0.0000
display	-0.0285	0.0058	0.0000
electronic	0.0315	0.0068	0.0000
image	-0.0248	0.0057	0.0000
information	-0.0039	0.0075	0.6071
interface	0.0166	0.0094	0.0764
memory	-0.0222	0.0060	0.0002
mobile	-0.0032	0.0119	0.7866
system	-0.0143	0.0052	0.0060
text	-0.0124	0.0111	0.2627
user	0.0078	0.0059	0.1903
video	0.0272	0.0069	0.0001

'Audio'키워드에 유의한 영향을 미치는 키워드는 digital, display, electronic, image, memory, system, video 로 나타났다. 이를 통하여 audio와 관련된 애플의 기술은 이들 7개의 키워드에 기반한 기술에 의존하고 있음을 알 수 있다. 다음 표는 애플의 특허에 포함된 'Video'키워드를 반응변수로 하고 나머지 키워드들을 설명변수로 한 베타회귀분석 결과이다.

유의확률이 0.05보다 작아서 'Video'키워드를 통계적으로 의미 있게 설명해 주는 키워드는 audio, device, digital, display, electronic, image, system, text로 모두 8개의 기술 키워드로 구성됨을 알 수 있었다. 이를 통하여 애플의 'Video'기술은 이들 키워드에 의존하는 하위 기술에 영향을 받고 있음을 알 수 있다. 다음은 마지막으로 'Text'특허 키워드를 반응변수로 하는 베타회귀모형의 분석결과를 나타내고 있다.

총 9개의 기술 키워드인 computer, device, digital, display, interface, memory, system, user, video 가 'Text'키워드에 통계적으로 유의한 영향을 미치고 있음이 확인 되었다. 본 연구에서는 'Data', 'Audio', 'Video', 그리고 'Text'의 4개 특허 키워드를 중심으로 나머지

표 3. 베타회귀분석 결과: 반응변수=video

Table 3. Beta regression result: response variable=video

Keyword	Estimate	S.E.	p-value
audio	0.0253	0.0066	0.0001
computer	-0.0006	0.0071	0.9367
content	-0.0025	0.0090	0.7778
device	-0.0194	0.0048	0.0001
digital	0.0533	0.0099	0.0000
display	0.0431	0.0058	0.0000
electronic	-0.0205	0.0067	0.0021
image	0.0487	0.0059	0.0000
information	-0.0035	0.0075	0.6373
interface	-0.0050	0.0092	0.5873
memory	-0.0044	0.0060	0.4647
mobile	-0.0150	0.0117	0.2030
system	0.0268	0.0052	0.0000
text	-0.0266	0.0110	0.0155
user	-0.0086	0.0059	0.1433

표 4. 베타회귀분석 결과: 반응변수=text

Table 4. Beta regression result: response variable=text

Keyword	Estimate	S.E.	p-value
audio	-0.0047	0.0042	0.2649
computer	0.0258	0.0046	0.0000
content	0.0006	0.0058	0.9196
device	-0.0066	0.0031	0.0341
digital	-0.0177	0.0062	0.0045
display	0.0105	0.0037	0.0046
electronic	-0.0023	0.0043	0.6009
image	-0.0022	0.0036	0.5523
information	-0.0038	0.0048	0.4330
interface	-0.0205	0.0060	0.0006
memory	-0.0082	0.0039	0.0334
mobile	-0.0074	0.0076	0.3310
system	0.0088	0.0033	0.0084
user	0.0298	0.0038	0.0000
video	-0.0107	0.0043	0.0136

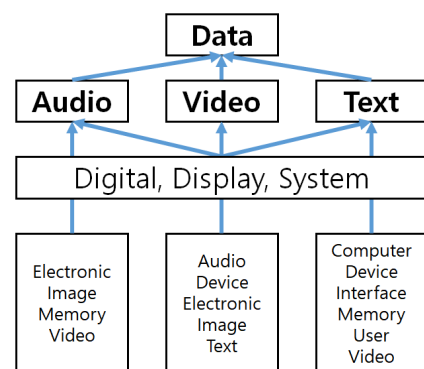


그림 3. 애플의 기술 연관 계층도

Fig. 3. Hierarchical diagram of technology relation for Apple

키워드들의 기술 영향력을 평가하였다. 표 1에서 표 4까지의 결과를 이용하여 그림 3에서 나타난 애플의 기술 연관 계층도를 구축하였다. 가장 상위에 있는 'Data'기술은 'Audio', 'Video', 그리고 'Text'

기술에 영향을 받고 있으며 또 'Audio', 'Video', 'Text'의 각 기술은 'Digital', 'Display', 그리고 'System'에 기반한 세부기술에 영향을 받고 있음을 알 수 있다. 그리고 'Audio', 'Video', 'Text'은 각각 그림 3에서 하위에 연결된 키워드들과 연관성이 있음을 알 수 있다.

4. 결론 및 향후 연구

본 논문은 R 데이터 언어를 이용한 특허문서의 전처리와 베타회귀분석을 이용한 특허 마이닝에 대하여 연구하였다. 개발된 기술에 대한 방대한 결과를 포함하고 있는 특허 문서를 분석하여 기업의 연구개발 전략을 위한 지식을 제공하는 방법에 대한 연구 필요성이 꾸준히 제기되어 왔다. 특히 기술은 시간에 따라 빠르게 진화하는 성장곡선을 따르기 때문에 이에 대한 고려가 있어야 한다. 베타회귀분석 모형의 반응변수는 0에서 1사이의 값을 갖고 이는 성장곡선의 특성을 포함하고 있다. 따라서 본 연구에서 특허 마이닝 방법은 반응변수의 성장성을 반영한 분석결과를 제공하여 보다 효율적인 기술경영을 위한 결과를 제공한다. 시간에 따른 특허 키워드의 움직임을 보다 정교하게 모형화하기 위하여 베타회귀분석 뿐만 아니라 FDA(functional data analysis)등의 고려를 통한 특허 마이닝 방법에 대한 연구가 진행 될 예정이다.

References

- [1] A. T. Roper, S. W. Cunningham, A. L. Porter, T. W. Mason, F. A. Rossini, and J. Banks, *Forecasting and Management of Technology*, Hoboken, NJ, John Wiley & Sons, 2011.
- [2] J. Keller, and H. A. V. D. Gracht, "The influence of information and communication technology (ICT) on future foresight processes - Results from a Delphi survey", *Technological Forecasting and Social Change*, Vol. 85, pp. 81-92, 2014.
- [3] J. Kim, J. Lee, S. Park, and D. Jang, "Technology Strategy based on Patent analysis", *Journal of Korean Institute of Intelligent Systems*, Vol. 26, No. 2, pp. 141-146, 2016.
- [4] S. Jun, "Technology Forecasting using Bayesian Discrete Model", *Journal of Korean Institute of Intelligent Systems*, Vol. 27, No. 2, pp. 179-186, 2017.
- [5] J. Kang, J. Kim, J. Lee, S. Park, and D. Jang, "Methodology of Prior Art Search Based on Hierarchical Citation Analysis", *Journal of Korean Institute of Intelligent Systems*, Vol. 27, No. 1, pp. 72-78, 2017.

- [6] S. Jun, "A Big Data Preprocessing using Statistical Text Mining", *Journal of Korean Institute of Intelligent Systems*, Vol. 25, No. 5, pp. 470-476, 2015.
- [7] D. Hunt, L. Nguyen, and M. Rodgers, *Patent Searching Tools & Techniques*, Hoboken, NJ, Wiley, 2007.
- [8] Y. Tseng, C. Lin, and Y. Lin, "Text mining techniques for patent analysis", *Information Processing and Management*, Vol. 43, No. 5, pp. 1216-1247, 2007.
- [9] R Development Core Team, *R: A language and environment for statistical computing*, R Foundation for Statistical Computing, Vienna, Austria, <http://www.R-project.org>, 2018.
- [10] I. Feinerer, K. Hornik, and D. Meyer, "Text mining infrastructure in R", *Journal of Statistical Software*, Vol. 25, No. 5, pp. 1-54, 2008.
- [11] I. Feinerer, and K. Hornik, *Package 'tm' Ver. 0.7-4*, Text Mining Package, CRAN of R project, 2018.
- [12] R. V. Hogg, J. M. McKean, and A. T. Craig, *Introduction to Mathematical Statistics*, 8th edition, Upper Saddle River, NJ, Pearson, 2018.
- [13] S. L. P. Ferrari, and F. Cribari-Neto, "Beta Regression for Modelling Rates and Proportions", *Journal of Applied Statistics*, Vol. 31, No. 7, pp. 799-815, 2004.
- [14] F. Cribari-Neto, and A. Zeileis, "Beta Regression in R", *Journal of Statistical Software*, Vol. 34, No. 2, pp. 1-24, 2010.

저자 소개



전성해(Sunghae Jun)

1993년 : 인하대학교 통계학과 이학사
 1996년 : 인하대학교 통계학과 이학석사
 2001년 : 인하대학교 통계학과 이학박사
 2007년 : 서강대학교 컴퓨터공학과 공학박사
 2013년 : 고려대학교 정보경영공학과 공학박사
 2003년~현재 : 청주대학교소프트웨어융합학부빅데이터통계학전공 교수

관심분야 : Artificial Intelligence, Statistical Learning, Big Data

Phone : +82-43-229-8205

E-mail : shjun@cju.ac.kr