

# 빅데이터를 이용한 자동 이슈 분석 시스템

## An Automatic Issues Analysis System using Big-data

최동열\*, 안은영\*\*

대덕대학교 자동차학과\*, 한밭대학교 정보통신공학과\*\*

Dongyeol Choi(dychoi@ddu.ac.kr)\*, Eungyoung Ahn(aey@hanbat.ac.kr)\*\*

### 요약

빠르게 변화하는 온라인상의 정보 흐름과 트렌드를 이해하고 IT기술 환경변화에 대응하기 위해서 필요한 선제적 제도 마련을 위한 한 가지 방안으로 빅데이터를 이용하고자 하는 노력이 최근 들어 더욱 가속화 되고 있다. 논문에서는 인공지능 기반의 빅데이터 처리를 통한 이슈 분석 시스템의 개발과 연구를 통해 빅데이터 처리를 위한 새로운 기술의 가능성을 확인하고자 한다. 이를 위해, 고속의 병렬처리가 가능해진 인공지능망을 사용, 의미 추론 및 패턴분석을 위한 처리 기법을 제안하고 구현을 통해 제안하는 방법에 대한 빅데이터 처리의 적합성을 알아본다. 정보보안의 중요성을 감안하여, 인공 신경망을 이용한 이슈 분석 시스템을 최근의 보안 이슈 분석에 활용해봄으로써 제안하는 방식이 실제 빅데이터 처리에 유용하게 활용 될 수 있음을 검증한다. 실험을 통해서 제안된 방식에 대한 다양한 목적의 빅데이터 처리를 위한 기반 기술로의 활용 가능성을 확인한다.

■ 중심어 : 빅데이터 | 사회관계망 | 인공지능망 | 구문 패턴 | 이슈 분석 |

### Abstract

There have been many efforts to understand the trends of IT environments that have been rapidly changed. In a view point of management, it needs to prepare the social systems in advance by using Big-data these days. This research is for the implementation of Issue Analysis System for the Big-data based on Artificial Intelligence. This paper aims to confirm the possibility of new technology for Big-data processing through the proposed Issue Analysis System using. We propose a technique for semantic reasoning and pattern analysis based on the AI and show the proposed method is feasible to handle the Big-data. We want to verify that the proposed method can be useful in dealing with Big-data by applying latest security issues into the system. The experiments show the potentials for the proposed method to use it as a base technology for dealing with Big-data for various purposes.

■ keyword : Big-data | Social Networks | Artificial Neural Networks | Syntax Pattern | Issue Analysis |

## I. 서 론

빅데이터를 이용한 기계학습 및 딥러닝에 있어서 각 응용 분야별로 학습에 필요한 데이터의 수집과 더불어

학습의 입력패턴을 어떻게 모델링 할 것인가가 매우 중요한 문제가 될 수 있다. 특히 빅데이터의 비정형성, 실시간적 성격을 감안한다면 더욱 그러하다.

어떤 단어들의 집합이 그 보다 상위의 의미를 갖는

\* 본 연구는 한밭대학교 교내연구비 지원을 받아 수행됨.

접수일자 : 2019년 12월 17일

수정일자 : 2020년 01월 15일

심사완료일 : 2020년 01월 15일

교신저자 : 안은영, e-mail : aey@hanbat.ac.kr

단어로 추상화될 수 있는데 본 논문에서는 이러한 추상화된 상위 단계의 단어를 이슈(issue)라고 정의한다. 이슈는 실시간으로 생성되는 데이터로부터 확인하여 추천시스템이나 의사결정 시스템 등의 주요 입력 값으로 활용될 수 있다. 특정 이슈를 구성하는 어휘들은 자연어로서 불규칙한 특성을 갖는다. 따라서 추상적 의미의 이슈를 구성하는 어휘의 군집이 다소 불완전하더라도 이를 인식할 수 있도록 학습의 입력패턴이 설계되어야 한다. 최근에는 자연어 처리 분야에서도 딥러닝을 이용하여 성능을 향상시키고자 하는 연구가 지속적으로 진행되고 있다[1][2]. 앞서 제안된 방법들은 토큰화 방법의 선정이나 토큰의 개수 선정에 따라 성능에 영향을 받는다. 또한, 특정 이슈들은 단어의 구성이나 개수에서 자연어로서 불규칙한 특성을 갖기 때문에 추상적 의미의 이슈를 구성하는 어휘의 군집이 불완전하더라도 이를 분류할 수 있도록 학습의 입력패턴이 설계되어야 하는데 이러한 문제들을 여전히 해결하고 있지 못하다. 본 연구에서는 온라인상의 실시간 빅데이터에 대한 이슈 분석 시스템을 개발하여 일련의 어휘들로부터 추상화된 상위개념의 의미를 찾아냄으로써 현재의 이슈 및 동향을 파악하는 것을 주된 목적으로 하며, 이를 위해 자연어의 특성이 갖는 이러한 비정형적 데이터에 대한 학습 입력패턴 생성 방안을 제시한다.

## II. 관련연구

모든 영역의 데이터가 범람하는 빅데이터 시대의 도래로 인해 최근 들어 인공지능 기술의 중요도가 부각되고 있다. 이 장에서는 빅데이터 지식처리를 위한 인공지능기술의 국내 연구사례를 간략하게 살펴본다. 2013년부터 10년간 인간의 지식증강서비스를 위해 빅데이터로부터 스스로 학습하여 지식을 축적하고, 시스템 및 기기 간의 자율협업방식으로 새로운 문제를 해결하려는 연구가 진행되고 있다. 그 중 하나인 지능 진화형 질의응답 시스템인 WiseQA의 개발은 자연어 질문을 이해하고, 비정형 및 정형 지식베이스를 통해 정답 후보들을 추론하고 순위화, 필터링 등의 과정을 거쳐 최종 정답을 근거와 함께 추론하는데 필요한 모듈을 개발

한다[3]. 이 시스템에서는 의미정답유형 인식을 위해 다중클래스 분류에 기반한 기계학습모델을 사용한다[4]. 빅데이터에 대한 자연어의 이해는 본 연구와 연관성이 있는 분야로서, 본 연구는 인간이 복잡한 일련의 현상을 어떻게 보다 포괄적 의미로 묶고 추상화를 통해서 패턴화하는가에 대한 인간의 이해 과정에 관한 것과 깊이 관련이 있다.

인공지능을 구현하는 구체적 접근방식을 살펴보면, 기계학습은 알고리즘을 이용해 데이터를 분석하고, 분석을 통해 학습하며, 학습한 내용을 기반으로 판단이나 예측을 하며 컴퓨터 비전 등의 분야에서 높은 인식률을 보이지만 학습된 이미지 일부의 훼손에도 인식률이 크게 떨어지는 단점이 있다. 따라서 기계학습을 사용하는 [5]의 논문에서도 이러한 한계를 해결하기 위한 방안으로 하이브리드 방식을 제안하기도 하였다[4]. 한편 딥러닝은 완전한 머신러닝을 실현하기 위한 기술로, 뇌의 뉴런과 유사한 정보 입출력 계층을 활용해 데이터를 학습하는 방법이며[6] 최근 병렬 연산에 최적화된 GPU가 등장함으로 인해 학습을 위한 대량의 연산과 속도 문제를 해결할 수 있게 됨으로써 여러 분야에서 연구가 이루어지고 있다[7][8]. 본 연구는 산업계와 학계에서 현재 진행 중인 빅데이터의 지식처리를 위한 자연어의 의미추론에 관한 연구로서 그 대안을 제시하고자 한다. 이를 위해 인공지능망 모델을 이용하여 인터넷상에서 유통되는 대량의 데이터로부터 현재의 경향과 이슈를 확인하는 분석시스템을 제안한다.

## III. 제안방법

본 연구에서 제안하는 이슈 분석 시스템의 구조는 [그림 1]과 같다. [그림 1]의 '기타 정보 수집'은 특정 IP 주소나 '특정 데이터 패턴' 등의 수집 기능 등 단순 어휘가 아닌 이슈 이미지로의 매핑이 가능한 정보의 수집을 말한다. 인공지능망을 이용한 빅데이터 이슈 분석 시스템을 위해서는 수집되는 어휘들을 신경망 회로의 입력으로 사용할 수 있도록 가공하여야 한다. 이에, 본 장에서는 인공지능망을 위한 이슈 이미지의 형상화 방안을 제안하고자 한다. 또한 특정 이슈에 대한 어휘의 시간

적 지속성을 감안하기 위해 네오 코르네티론에 시간 지속성에 관한 속성인 TLWS(Time Length Work Sheet)을 가진 기억소멸형 플레인(decay memory plane)을 제안하고 그 개념과 필요성에 대해 설명한다. 마지막으로 이를 적용한 이슈 분석 시스템에 대해 논의한다.

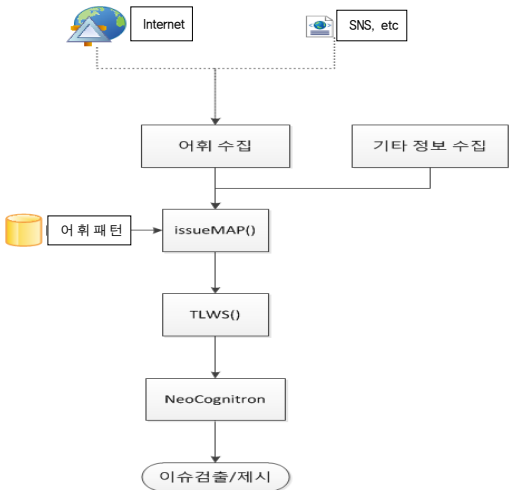


그림 1. 이슈 분석 시스템 개요도

## 1. 이슈 패턴의 생성

### 1.1 어휘별 입력 패턴 생성규칙

몇 개의 관련 어휘들을 포괄적 의미로 묶고 추상화하여 이해가 가능한 수준으로 추론하는 이슈 분석 시스템을 신경망 회로를 이용하여 구현하기 위해서는 먼저 수집되는 어휘들을 신경망 회로의 입력으로 사용할 수 있도록 가공하여야 한다. 이를 위해서는 어휘에 대한 한글 문자 자체를 패턴화하거나 어휘의 의미에 따른 이미지를 사용하는 방법 등 다양한 방법을 사용할 수 있다. 추론을 위한 입력패턴으로 사용되는 어휘는 명사와 동사가 된다. 신경망 회로의 입력으로 사용할 수 있는 어휘들에 대한 패턴화 방법이 다양하게 제시될 수 있으나 본 논문에서는 이슈를 이미지 패턴화하고 이슈를 정의하는 어휘들을 이슈 이미지의 서브 이미지패턴으로 관계화 하는 방법을 제안한다.

[그림 2]에서 '학습된 실제 의미 패턴'은 각 의미마다 부여되는 이미지다. 어휘1-1이란 이슈\_1를 구성하는

어휘 중 1번째 어휘라는 의미이다. 부분패턴 (.), T, L, T 등은 각 어휘가 이미지의 어느 부분에 관계되어 있는지를 표시한 것이다. 특별한 의미를 부여해서 관계를 지을 수도 있지만 반드시 그럴 필요는 없다. 이슈맵은 특정 시점에 나타난 어휘를 가지고 이슈 이미지를 구성하는 것으로 [그림 2]에서는 T, (,)에 해당하는 세 개의 어휘가 관찰되었다는 의미이다.

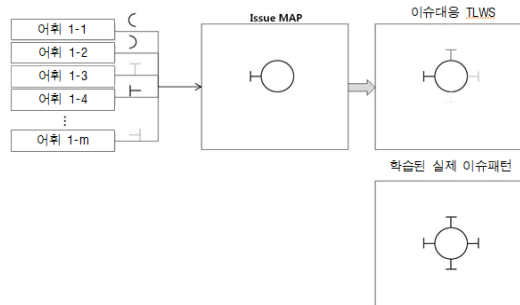


그림 2. 몇 개의 어휘들로 구성되는 의미 맵

### 1.2 의미 패턴 생성규칙

여기서는 3.1.1에서 제시한 이슈 이미지 패턴화 규칙을 사용하여 제안하는 의미 분석 시스템의 실험을 위한 실례로서 보안문제에 관한 의미 분석을 적용해보기로 한다. 이를 위해 보안문제와 관련한 의미 패턴 및 구성 어휘를 생성 해보기로 한다. [표 1]은 각 의미들에 대한 구성어휘들을 조합한 의미 패턴의 예들을 보여 주고 있는데 각 이슈 패턴들은 구성 어휘들이 나타내는 서브 이미지들을 모두 합성한 이미지가 된다.

표 1. 이슈 패턴과 그 구성 어휘의 서브 이미지

이슈	이슈 이미지	구성 어휘	서브 이미지	클래스 #
모바일 사이버 범죄 등장과 스마트폰기 타깃 다각화		스미싱		0 번
		T2W		
		스마트폰		
		결재		



[표 1]에서 보여주는 바와 같이 특정 어휘가 검색되면, 의미맵 생성 모듈은 그 어휘에 해당하는 서브 이미지를 중첩, 합성하여 출력한다. 여러 개의 관련 어휘가 검색되면 그에 따른 여러 개의 서브 이미지가 출력 이미지에 합성된다. 출력 이미지에 합성될 때에는 다음 절에서 설명하는 TLWS의 공식에 따라서 합성되어 어휘가 검색되는 빈도와 주기에 따른 정보를 적용하게 된다.

2. TLWS 함수와 기억소멸형 플레인(plane)

데이터 수집은 통상적으로 모든 데이터를 수집하고 처리할 수 없으므로 순서에는 관계 없지만 직렬처리의 형태로 이루어지게 된다. 특정 이슈와 연관된 어휘들은 임의의 시간 간격을 두고 나타날 수 있다. 시간의 관점에서, 짧은 시간 안에 나타나는 어휘들은 동시에 나타난 어휘들이라고 간주 할 수 있다. 또한 현재 시점에서 보았을 때 나타나지 않는 어휘들은 해당 이슈에 대한 효력이 시간이 경과됨에 따라 약해져 간다고 볼 수 있다. 이러한 관계를 정의하기 위해 TLWS함수(TLWS Activation Function)를 설계하고 기억소멸형 플레인(Decay Memory Plane)을 인공신경망에 추가적함으로써 어떠한 의미를 구성하는 어휘들의 출현 시간과 빈도를 고려한 의미추론이 가능하도록 한다. TLWS함수는 특정 어휘의 출현 빈도수가 높으면 1에 가까워지는 출력을 내고, 어휘의 출현 이후 시간이 지날수록 쇠퇴해서 결국 0에 수렴하는 출력을 내주는 역할을 셀이면서 1로 표시된다. 식 1에서 R은 입력 의미맵 이미지의 단위 값으로 [1, 0]이며,  $a'(x)$ 는 바로 전 시점의  $a(x)$ 값이다.

$$\begin{aligned}
 a(x) &= k'(x) - e(x) - v(x) \\
 \text{where, } e(x) &= (1 - a'(x)) \times 0.25R, \\
 v(x) &= 0.125a'(x) \times (1 - R)
 \end{aligned}
 \tag{1}$$

어떤 어휘가 특정 시점에 관찰되면 이슈 이미지에서

그 어휘에 대한 서브 이미지가 TLWS함수에 따라서 활성화된다. 또한 특정 시점에 나타나지 않은 어휘에 해당하는 서브 이미지는 TLWS함수에 따라서 쇠퇴하게 된다. 이렇게 생성된 이슈 대응 TLWS 출력은 모든 이슈에 대해 학습을 완료한 인공 신경망의 입력 값으로 순차적으로 입력된다. TLWS 출력에 따라서 인공신경망은 해당 이슈가 활성화 되는지를 파악하여 출력하는데 어휘들은 여러 이슈에 복합적으로 관여할 수 있으므로 특정 시점에는 여러 개의 이슈가 파악될 수도 있으며 해당 이슈의 가능성을 [0.0 ~ 1.0]의 실수로 출력한다.

3. 이슈 분석을 위한 학습망 구조

인터넷에서 수집되는 방대한 양의 어휘 데이터들을 감안하여 사전에 정의된 이슈에 관계된 어휘들을 추려내는 전처리 과정을 거치게 된다. 전처리과정은 텍스트 마이닝 분석 기법의 텍스트 마이닝 과정 등을 통해 사전에 정의된 이슈에 관계된 어휘들을 추려내는 것이 가능하다. [그림 3]은 빅데이터 처리를 위한 이슈 분석 시스템 구조를 설명하기 위한 것으로, 이슈 맵(Issue MAP)을 고안하여 문자 정보인 어휘를 신경망의 입력인 이미지 패턴 형식으로 바꿀 수 있도록 하고 기억 소멸형 플레인 TLWS함수를 인공 신경망에 접목하는 방법으로 시간차를 두고 나타나는 어휘들로부터 특정 이슈를 파악할 수 있도록 하였다.

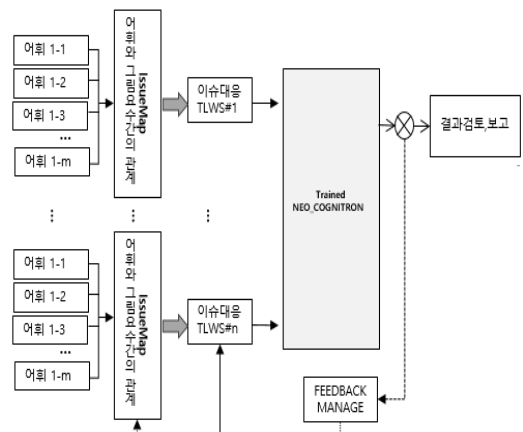


그림 3. 이슈 분석 시스템의 구조

[그림 3]은 인공 신경망에 입력되기까지 어휘 수집, 이슈맵 및 기억 소멸형 플레인을 거치는 데이터흐름의 전과정을 보여준다. 여러 경로로 수집되는 어휘는 이슈맵 생성 함수로 전달된다. 이 때, 각 어휘는 해시테이블을 통해 연관된 패턴 이미지의 데이터베이스에서의 위치를 구하고, 이렇게 찾아진 패턴 이미지들은 논리적으로 합해져서 이슈 이미지로 만들어지게 된다. 만들어진 이슈 이미지는 TLWS 함수에 전달되어 TLWS 연산과 식 2의 임계값(threshold) 연산을 거친 후 인공신경망의 첫 번째 단계인 UO(Retina)에 입력된다.

$$I_{noise} = \begin{cases} I_{max} & \text{if } P \geq \text{Threshold} \\ 0 & \text{otherwise} \end{cases} \quad P \in [0.0, 1.0] \quad (2)$$

#### IV. 실험 및 결과

##### 1. 실험 환경 및 조건

3.3절에서 제안한 빅데이터 이슈 분석 시스템의 보안 이슈 적용에 대한 실험을 위해 실험 환경은 [표 2]와 같다. 본 실험에는 총 6가지의 보안 이슈를 가정하고 각 이슈는 5개 이상의 어휘로 구성되도록 하였다[9][10]. 이슈 및 어휘 구성은 [표 3]에 표시된 바와 같다. 학습에는 총 7개의 '이슈 이미지'를 사용하였는데 이들 중 2개의 '이슈 이미지'는 같은 이슈를 나타내는 것으로 해서 총 6개의 이슈를 검출하도록 하였다.

표 2. 실험환경

시스템 요소	구성
하드웨어	데스크톱 개인용 컴퓨터
운영체제	마이크로소프트 윈도우즈 7.
인공신경망	MLP(Python 버전)

[표 3]의 첫 번째 이슈(이슈 클래스 번호 0번)는 [표 1]에 보인 바와 같은 6개의 어휘로 구성되는데, 어떤 시점에 모든 어휘가 다 나타나면 완전한 형태의 이슈 맵이 생성되지만 그렇지 않은 경우에는 일부의 패턴으로 이슈 맵을 형성하게 된다.

표 3. 학습데이터와 구성어휘

이슈	이슈 패턴	구성 어휘	클래스 #	비고
모바일 사이버 범죄등장과 스마트기기 타깃 다각화		스미싱, T2W, 스마트폰, 결제, 암호, 보안카드	0	
중비 스마트폰, 사이버테러에 본격화		주소록, 위치추적, 원격제어, 무선, 제작물	1	
변칙적 광고성 프로그램의 증가와 복잡화		프로그램, 포털사이트, 마케팅, 광고, Trojan horse	2	
업데이트 프로그램 변조를 통한 악성화일 지능화		DDos, 사이버전, 업데이트, 변조, 소프트웨어	3	
문서 형 악성화일의 은밀한 표적공격의 고도화		HWP, 문서화일, 취약점, 표적공격, 한컴오피스	4	
전자금융사기와 사이버 범죄의 산업화		넷버스, 악성앱, 승인문자, 소액결제	5	
문서 형 악성화일의 표적공격		한국, 맞춤형, 문서파일, 오피스	6	4와 유사

##### 2. 성능 분석

본 연구에서 제안하는 이슈 분석 시스템의 성능 분석을 위해 인공신경망에 7개 이슈에 대한 이슈 이미지의 학습을 진행하고, 총 7개의 이슈에 대해 각 이슈별로 4가지 상황을 가정하여 총 28개의 입력 패턴을 처리하도록 하였다. [표 4]와 같은 입력패턴에서 실험을 진행하여 이슈 분석 시스템이 입력 패턴에 대해 해당 이슈를 분석한 결과를 확인해보았다.

이슈별 입력 패턴에 대해 얼마나 정확하게 해당 이슈를 분리해 내는지에 대한 측정하기 위해 분리도를 확인하였다. [표 5]는 실험에 사용된 이슈별 패턴에 대한 분리도를 측정된 결과이다. 특정 입력 패턴에 대한 이슈 분석 결과 값을  $f(x)$ 라 하고, 다른 클래스로 나타난 결과 값 중에서 가장 큰 값을  $g(x)$ 라고 할 때  $f(x)$ 가 0 이상 일 때 분리도  $q(x)$ 를 식 3과 같이 정의한다.

$$q(x) = \left[ 1.0 - \frac{g(x)}{f(x)} \right] \quad (3)$$

$$q(x) = \begin{cases} 0.0 & \text{if } f(x) \equiv 0, \\ 1.0 & \text{if } g(x) \equiv 0 \end{cases}$$

$$x = \begin{cases} 1.0 & \text{if } x > 1.0 \\ x & \text{else} \end{cases}$$

표 4. 실험 데이터

이슈 0		is0-0	이슈 1		is1-0
		is0-1			is1-1
		is0-2			is1-2
		is0-3			is1-3
이슈 2		is2-0	이슈 3		is3-0
		is2-1			is3-1
		is2-2			is3-2
		is2-3			is3-3
이슈 4		is4-0	이슈 5		is5-0
		is4-1			is5-1
		is4-2			is5-2
		is4-3			is5-3
이슈 6		is6-0			
		is6-1			
		is6-2			
		is6-3			

[표 5]의 결과를 활용하여 이슈 분석 시스템의 성능을 측정해보기로 한다. [표 5]에서 보듯이 이슈 0과 6은 이슈는 다르지만 유사한 이슈 이미지를 사용하였기 때문에 이슈 0과 이슈 6의 이슈 분석 성능 실험 결과를 보면 8개의 입력 패턴에 대해 두 개 이슈 0과 6에서 동시에 출력 값이 나타나고 있다. 이 결과를 각 각의 분석 성능 백분율과 분리도 그래프로 나타내보면 제안된 방식의 이슈 분석 성능을 확인해 볼 수 있다. [그림 4][그림 5]의 분석 성능 백분율과 분리도 그래프를 보면 이슈 0의 결과를 얻을 때 비슷한 이미지를 사용한 이슈 6의 결과도 일부 도출되었으나 이슈 0에 해당하는 입력패턴에 대해서는 이슈 0의 결과가 확연하게 더 높은 비율을 나타내고 있으며 분리도 값도 높게 나타난다.

또한 이슈 6의 결과를 얻을 때 비슷한 이미지를 사용한 이슈 0의 결과도 일부 도출되었으나 이슈 6에 해당하는 입력 패턴에 대해서는 이슈 6의 결과가 이슈 0의 결과보다 더 높은 비율을 나타내는 것을 볼 수 있다. 분리도 측면에서 보면 이슈 6의 이미지가 이슈 이미지 0과 유사하여 이슈 이미지 분리도가 낮게 나타는 현상을 보이고 있기는 하지만 서로 다른 이슈에 대해 이슈를 구성하는 어휘가 비슷하게 구성되었더라도 정확하게 해당 이슈를 분석 해내고 있음을 보이고 있다.

표 5. 실험데이터에 대한 분리도 측정 결과

이슈	입력 패턴	분리도		
		f(x)	g(x)	q(x)
이슈 0	is0-0	0.4202	0.0929	0.7789
	is0-1	0.3164	0.0154	0.9513
	is0-2	0.1423	0.0000	1.0000
	is0-3	0.3741	0.1490	0.6017
이슈 1	is1-0	0.2213	0.0000	1.0000
	is1-1	0.0268	0.0000	1.0000
	is1-2	0.2027	0.0000	1.0000
이슈 2	is1-3	0.2135	0.0000	1.0000
	is2-0	0.6011	0.0414	0.9311
	is2-1	0.5925	0.0857	0.8554
	is2-2	0.5816	0.0738	0.8731
이슈 3	is2-3	0.5779	0.0872	0.8491
	is3-0	0.6148	0.1328	0.7840
	is3-1	0.3352	0.0000	1.0000
이슈 4	is3-2	0.6105	0.4634	0.2410
	is3-3	0.3146	0.1848	0.4126
	is4-0	1.4576	0.0000	1.0000
	is4-1	1.2054	0.0000	1.0000
이슈 5	is4-2	1.4611	0.0000	1.0000
	is4-3	1.4638	0.0000	1.0000
	is5-0	0.3622	0.0000	1.0000
	is5-1	0.2588	0.0000	1.0000
이슈 6	is5-2	0.2597	0.0000	1.0000
	is5-3	0.1491	0.0000	1.0000
	is6-0	0.4001	0.2593	0.3519
	is6-1	0.2568	0.1862	0.2749
	is6-2	0.3116	0.2592	0.1682
	is6-3	0.4009	0.2612	0.3485

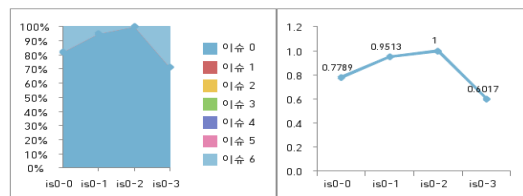


그림 4. 이슈 0 입력패턴 분석 성능 백분율과 분리도

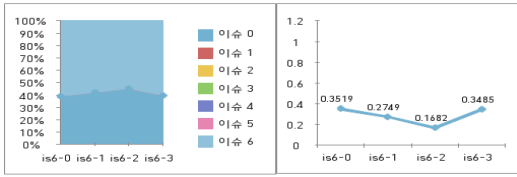


그림 5. 이슈 6 입력패턴 분석 성능 백분율과 분리도

## V. 결론

인터넷 환경에서 검출될 수 있는 문자 기반 빅데이터로부터 어휘를 검출하여 이로부터 특정 이슈를 파악하는 시스템을 인공지능을 이용해서 구현하고 보안과 관련한 이슈에 대해 제안된 방법을 적용하여 모의실험을 진행 하였다. 실험을 통한 성능 분석 결과 제안한 이슈 분석에 대한 성능의 우수성을 확인하였다. 제안된 방법은 이슈 이미지에 포함되는 중심어휘의 수를 제한하지 않기 때문에 기존의 심층학습을 통한 자연어 처리 방식에서 문제가 되는 학습패턴의 생성 방법의 한 가지 대안이 될 수 있다. 더욱이 시간소멸형 함수를 학습망의 앞 단계에 도입함으로써 이슈를 구성하는 중심 어휘들의 시간에 따른 사용 빈도수를 반영할 수 있다. 제안된 방법은 사회의 여러 방면에서 선제적인 정책 이슈 발굴과 기획을 위한 분석 시스템으로서의 기초적 역할을 할 수 있을 것이다. 그러나 향후, 본 논문에서는 논의 되지 못했던 이슈 맵의 기능을 확장하여 비정형 데이터 베이스를 활용하거나 IP주소 및 기타 패킷 정보를 활용할 수 있도록 하는 추가 연구가 필요하며 기억 소멸 형 플레인과 중간 단계의 셀들 값을 관찰하여 새로운 이슈의 등장을 모니터링하고 제어할 수 있는 추가적인 연구가 필요하다.

## 참 고 문 헌

- [1] 김도우, 구명완, "Doc2Vec과 Word2Vec을 활용한 Convolutional Neural Network 기반 한국어 신문 기사 분류," 정보과학회논문지, 제44권, 제7호, pp.742-747, 2017.
- [2] 이태겸, 신경섭, "심층 신경망 기반 자연어처리 모델의

성능 비교," 한국통신학회논문지, 제44권, 제7호, pp.344-1350, 2019.

- [3] 이형직, 류범모, 임수중, 장명길, 김현기, "빅데이터 지식 처리 인공지능 기술동향," ETRI 전자통신동향분석, 제29권, 제4호, pp.30-38, 2014.
- [4] 허정, 류범모, 김현기 외 1명, "WiseQA를 위한 정답 유형 인식," 정보처리학회논문지, 제4권, 제7호, pp.283-290, 2015.
- [5] Marius A. Pasca and Sandra M. Harabagiu, "High performance question/answering," Proceedings of the 24th annual international ACM SIGIR conference on Research and development in information retrieval, ACM, 2001.
- [6] K. FUKUSHIMA, "Cognitron: a self-organizing multi-layered neural network model," Biol. Cybern., Vol.20, pp.121-136, 1975.
- [7] 이택희, CPU-GPU 이기종 임베디드 시스템에서 딥러닝 추론의 가속, 서울시립대학교 전자전기컴퓨터공학과, 석사학위논문, 2018.
- [8] 구동균, Deep Learning을 이용한 택시 승객 승차 예측에 관한 연구, 서울시립대학교 교통공학과, 석사과정논문, 2018.
- [9] 김정덕, "국가 정보보안 이슈 및 정책방안에 관한 연구," 디지털정책 연구, 제10권, 제1호, pp.105-111, 2012.
- [10] ㈜시큐아이, 213년 상반기 10대 보인 이슈, 2013.

## 저 자 소 개

최 동 열(Dongyeol Choi)

정희원



- 1984년 2월 : 동국대학교 전자계산학과 학사
- 1987년 2월 : 동국대학교 전자계산학과 석사
- 2015년 2월 : 한밭대학교 멀티미디어공학과 박사
- 1993년 3월 ~ 2019년 2월 : 대덕대학교 컴퓨터공학과 교수

■ 2019년 3월 ~ 현재 : 대덕대학교 자동차학과 교수  
 <관심분야> : 인공지능, 자율주행, 빅데이터

안 은 영(Eungyoung Ahn)

중신회원



- 1989년 2월 : 동국대학교 전자계학과 학사
- 1991년 2월 : 동국대학교 컴퓨터공학 석사
- 2000년 8월 : 동국대학교 컴퓨터공학 박사
- 2000년 3월 ~ 2006년 3월 : 백석대학교 정보통신학부 조교수
- 2006년 4월 ~ 현재 : 한밭대학교 정보통신공학과 교수  
<관심분야> : 컴퓨터그래픽스, 가상현실, 시뮬레이터