

빅데이터 분석을 통한 아두이노 강의에 대한 사회적 인식

이은상
공주대학교

요약

이 연구의 목적은 빅데이터 분석 방법을 이용하여 아두이노 강의에 대한 사회적 인식을 분석하는 데 있다. 이를 위해 네이버 사이트의 블로그, 카페, 뉴스 채널에서 ‘아두이노+강의’를 검색 키워드로 2012년 1월부터 2021년 5월까지의 데이터를 텍스트 톰 사이트로 수집하였다. 수집된 데이터는 텍스트 톰 사이트를 이용하여 정제하였으며, 텍스트 톰 사이트, Ucinet 6, Netdraw 프로그램을 이용하여 텍스트 마이닝 분석과 의미 연결망 분석을 수행하였다. 빈도 분석, TF-IDF 분석, 연결 중심성 등의 텍스트 마이닝 분석 결과 ‘교육’, ‘코딩’ 등이 상위 키워드임을 확인하였다. 의미 연결망 분석을 위해 CONCOR 분석을 수행한 결과 ‘아두이노 관련 교육’, ‘피지컬 컴퓨팅 관련 강의’, ‘아두이노 특강’, ‘GUI 프로그래밍’ 등 4개의 군집을 확인할 수 있다. 이 연구를 통해 인터넷상에서 아두이노 강의와 관련하여 일반 대중들의 여러 가지 의미 있는 사회적 인식을 확인할 수 있었다. 이 연구의 결과는 아두이노 강의를 준비하는 교수자나 해당 주제를 연구하는 연구자, 나아가 소프트웨어 교육이나 코딩 교육과 관련 정책을 수립하는 정책 입안자들에게 의미 있는 시사점을 제공하는 자료로 활용될 것이다.

키워드 : 빅데이터 분석, 아두이노 강의, 의미 연결망 분석, 텍스트 톰, 텍스트 마이닝

Social perception of the Arduino lecture as seen in big data

Eunsang Lee
Kongju National University

Abstract

The purpose of this study is to analyze the social perception of Arduino lecture using big data analysis method. For this purpose, data from January 2012 to May 2021 were collected using the Textom website as a keyword searched for ‘arduino + lecture’ in blogs, cafes, and news channels of NAVER website. The collected data was refined using the Textom website, and text mining analysis and semantic network analysis were performed by opening the Textom website, Ucinet 6, and Netdraw programs. As a result of text mining analysis such as frequency analysis, TF-IDF analysis, and degree centrality it was confirmed that ‘education’ and ‘coding’ were the top keywords. As a result of CONCOR analysis for semantic network analysis, four clusters can be identified: ‘Arduino-related education’, ‘Physical computing-related lecture’, ‘Arduino special lecture’, and ‘GUI programming’. Through this study, it was possible to confirm various meaningful social perceptions of the general public in relation to Arduino lecture on the Internet. The results of this study will be used as data that provides meaningful implications for instructors preparing for Arduino lectures, researchers studying the subject, and policy makers who establish software education or coding education and related policies.

Keywords : Arduino lecture, big data analysis, semantic network analysis, text mining, TEXTOM

이 논문은 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(No. 2021R1F1A104755011).

논문투고 : 2021-10-16

논문심사 : 2021-11-16

심사완료 : 2021-11-21

1. 서론

오늘날 4차 산업혁명 시대를 맞이하여 과거에는 상상만으로 존재했던 다양한 기술들이 일상생활에서 속속 등장하고 있다. 3D 프린터, 드론, 자율 주행 자동차, AI, 사물인터넷 등의 기술은 불과 20여 년 전만 하더라도 영화 속에서도 소개되었지만, 오늘날에는 우리 생활 주변에서도 이들 기술을 쉽게 접할 수 있게 되었다. 이러한 시대적 흐름은 간단한 전자 제품을 만들어 보는 제작 활동에서도 혁신적인 변화를 일으켰다. 예를 들어, 과거에는 간단한 전자 제품을 만들기 위해 저항, 콘덴서, 다이오드, 트랜지스터, 브레드보드, 와이어, 각종 센서 등의 전자 부품에 대한 기본 지식이 필요했으며, 이들의 회로 구성에 대해서도 어느 정도 이해가 있어야 했다. 이들에 대한 지식이나 이해가 있다 하더라도 실제 전자 제품을 만드는 데 사용되는 저항 외의 전자 부품은 그 종류가 매우 많고 다양하기 때문에, 이들을 날개로 개별 구입하기 어려웠다. 이러한 어려움 때문에 고급 기능이 있는 전자 제품을 만드는 활동은 전자 분야에 특별한 취미가 있는 일부 마니아들이나 수행하는 활동으로 인식돼 왔다.

그러나 오늘날에는 이러한 전문적인 지식을 갖지 않아도 간단한 전자 제품을 만들 수 있는 시대가 열리게 되었다. 즉, 전자 부품에 들어가는 수많은 부품의 종류와 복잡한 회로 구성 방법을 익히지 않아도 과거보다 쉽고 빠른 방법으로 전자 제품을 제작할 수 있게 된 것이다. 특히, 오늘날에는 과거에는 구현할 수 없었던 사물과의 상호 작용이나 IoT 기능까지 구현할 수 있는데, 이와 같이 이 시대를 변화시킨 주인공은 바로 아두이노였다.

아두이노의 기원은 와이어링 보드에서 유래하였다. 와이어링 보드는 2003년부터 개발되어 2004년에 최초 버전이 나온 보드로, 당시 세계적으로 대중에게 잘 알려진 AVR 마이크로컴퓨터 보드였다. 이 보드는 통합개발 환경과 다양한 라이브러리를 제공함으로써 프로그래밍을 쉽게 할 수 있는 특징이 있었다. 와이어링 보드는 사용하기 쉬운 보드였지만, 비교적 고가의 부품을 사용했기에 보드 자체의 가격이 비싸 교육용으로 사용하기에는 무리가 있었다[23]. 이러한 문제점을 확인한 이탈리아의 마시모 벤지 교수팀은 와이어링 보드를 최대한

간략하게 구성하고 가격을 낮춘 보드를 개발하였는데, 이 보드가 바로 아두이노 보드였다[13].

아두이노의 외형은 일상생활에서 쉽게 볼 수 있는 컴퓨터 메인보드를 단순화하여 작게 만든 기판처럼 생겼는데, 이 기판에 여러 가지 센서나 부품을 쉽게 연결할 수 있다. 아두이노는 이와 같은 센서나 부품을 이용하여 외부의 정보를 수집하거나 필요한 기능을 수행할 수 있다. 즉, 다수의 스위치나 센서로부터 값을 입력받을 수 있고, LED, 릴레이, 모터와 같은 외부 전자 장치를 통제할 수 있는데, 이를 통해 과학 실험이나 사물인터넷같이 주변 환경과 상호 작용이 가능한 물건을 만들어 낼 수 있다[9]. 본래 이와 같은 기능은 마이크로컨트롤러 제어 보드에 의해 수행되는데, 아두이노의 가장 큰 장점은 이러한 보드에 대한 배경 지식이 없는 일반인이나 학생들도 쉽게 적용하고 이용할 수 있도록 개발되었다는 점이다[26]. 기존 마이크로컨트롤러는 특정 프로그램에서 소스 코드를 컴파일한 후 별도의 장치를 통해 업로드 해야 하는 복잡한 과정을 거쳐야 했지만, 아두이노는 무료 공개 프로그램을 이용하여 간단 방법으로 소스 코드를 컴파일하여 USB로 업로드 할 수 있었다[9]. 그 외에도 아두이노는 다른 마이크로컨트롤러 기반 보드보다 가격이 저렴한 점, 윈도우를 비롯한 다양한 운영체제에서 실행될 수 있는 점, 많은 실습 예제와 튜토리얼을 구할 수 있는 다양한 커뮤니티가 존재하는 점 등의 장점이 있었다. 이러한 장점으로 인해 엔지니어, 프로그래머, 미디어 아티스트, 디자이너뿐만 아니라 일반인에 이르기까지 다양한 사람들이 아두이노를 이용하게 되었다[18]. 특히, 최근 메이커 운동과 같은 문화적 현상에 힘입어 아두이노의 이용자는 점차 늘어나게 되었는데, 이에 따라 아두이노에 대한 교육 수요 역시 꾸준히 증가하게 되었다.

아두이노 관련 교육의 수요가 증가하였는지 확인할 수 있는 간단한 방법은 인터넷 포털 사이트나 유튜브 등의 온라인 플랫폼에서 해당 키워드를 검색해 보는 일이다. 예를 들어, 이들 온라인 플랫폼에서 ‘아두이노 강의’를 검색해 보면 수많은 사람이 만들어 놓은 콘텐츠를 확인할 수 있으며, 지금 이 순간에도 누군가에 의해 관련 콘텐츠를 제작되어 해당 내용이 공유되고 있다. 이러한 콘텐츠들은 개인의 자발적인 참여로 만들어진 것으로 이 시대를 살아가고 있는 대중의 사회적 인식을 그대로 반영하고 있다. 오늘날에는 컴퓨터 기능의 향상과

자료 분석 기술이 발달하여 이와 같이 온라인 플랫폼상에 쌓여 있는 많은 양의 자료를 분석할 수 있는 연구 방법이 보급되었는데, 이 방법이 바로 빅데이터 분석 방법이다. 빅데이터 분석 방법은 인간의 생활 속에서 일어나는 수많은 현상 및 변화와 관련된 대규모의 데이터를 이용하여 다양한 연관성을 분석하고 해석하는 데 사용되고 있으며[4], 특히, 특정 집단이 아닌 사회 전체적으로 나타나는 인식 및 현상을 파악하는 데도 활용되고 있다[1]. 또한, 이 분석 방법은 기존 연구 방법을 통해서 알기 어려웠던 사람들의 심리 및 행동 방식을 쉽게 파악하고 평가할 수 있어 경제, 정책, 문화, 교육, 마케팅, 스포츠 등 여러 분야에서 널리 사용되고 있다[16].

이 연구에서는 이러한 배경에 따라 빅데이터 분석 방법으로 아두이노 강의에 대한 사회적 인식을 분석하는 연구를 수행하고자 하였다. 이는 이미 학술적으로 아두이노 강의와 관련된 다양한 선행 연구가 수행된 적 있으나, 오늘날 대중의 삶을 반영하고 있는 온라인 플랫폼 상에서 축적된 많은 양의 아두이노 강의 관련 텍스트 자료를 분석한 연구는 수행된 적이 없기 때문이다. 이 연구의 결과는 아두이노 강의와 관련된 교수자나 연구자뿐만 아니라 최근 소프트웨어 교육과 같은 국가적 차원에서의 교육 정책을 추진하는 정책 입안자들에게 의미 있는 시사점을 제공하는 기초 자료가 될 것이다.

2. 연구 방법

2.1. 자료 수집

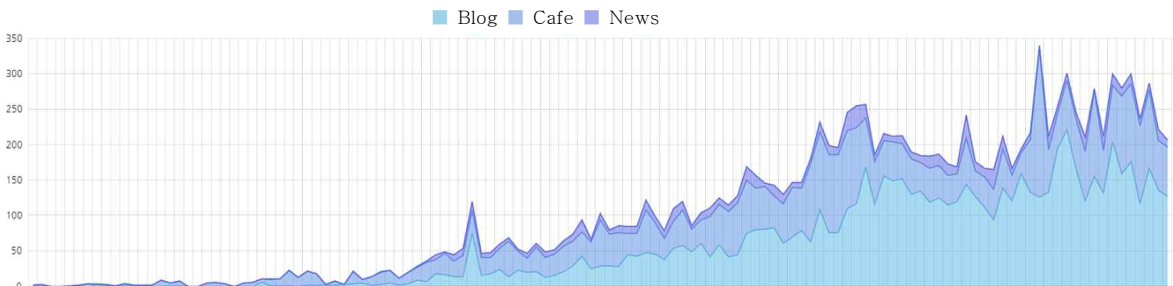
이 연구의 데이터 수집은 텍스트 사이트를 활용하였다. 이 사이트는 국내 유명 포털 사이트나 SNS상의 텍스트 데이터를 수집하여 데이터 세트(Data set)를 만들 수 있

는 사이트로, 단계적인 처리방식을 지원하여 효율적으로 빅데이터를 분석할 수 있었다[15]. 이 연구에서는 이 사이트에서 네이버의 블로그, 카페, 뉴스 등의 웹문서를 수집 채널로 선정하였으며, 검색 키워드는 ‘아두이노+강의’로 입력하였다. 네이버 사이트를 선정 한 이유는 같은 내용의 텍스트가 여러 포털 사이트에 동시에 노출된다는 점과 네이버가 국내 포털 사이트 중 80%의 사용 점유율을 차지하고 있는 점을 고려하였기 때문이다 [15, 21].

데이터의 수집 기간은 아두이노 관련 연구가 국내에서 시작된 시점인 2011년 1월부터 키워드 검색 당시 (2021년 5월)까지를 검색 기간으로 설정하였다. 그 결과 데이터의 수집량은 블로그 7,135건(2.24MB), 카페 4,997건(2.7MB), 뉴스 942건(350KB) 등으로 총 13,074건 (5.3MB)의 데이터가 수집되었다. 해당 기간별 ‘아두이노 + 강의’ 데이터의 수집량을 나타내면 (Fig. 1)과 같다.

2.2. 자료의 정제

자료의 정제는 총 2차에 걸쳐 진행되었다. 1차 정제는 수집된 13,074건의 데이터 중에서 1음절의 ‘위’, ‘등’, ‘것’, ‘월’ 등과 ‘관련’, ‘배우’, ‘오늘’, ‘안녕’ 등과 같이 특별한 의미가 없는 단어를 삭제하였다. 2차 정제는 1차 정제를 마친 단어 중에 유사어나 동의어를 통일하는 과정을 거쳤다. 예를 들어, ‘이용’, ‘사용’은 ‘활용’으로, ‘강좌’는 ‘강의’로 ‘아이’, ‘초등학생’, ‘중학생’, ‘고등학생’은 ‘학생’으로 변환하는 과정을 거쳤다. 자료의 정제는 Notepad ++ 프로그램을 이용하여 해당 키워드를 검색한 후 이를 삭제하거나 변환하는 작업을 수행하였다.



(Fig. 1) Data collection of 'Arduino + Lecture' by period

2.3. 자료의 분석 절차

이 연구에서 수행된 분석 절차는 크게 두 단계를 거쳐 진행하였다. 첫 번째 단계는 텍스트 마이닝 단계이다. 여기서 텍스트 마이닝이란 인간의 언어로 이루어진 대규모 비정형 텍스트 데이터로부터 연계성을 파악하거나, 분류, 군집화, 요약 등을 통해 관계 데이터를 추출하는 방식을 말한다[11]. 연구자는 텍스트 사이트를 이용하여 수집한 온라인 텍스트 데이터를 2차에 걸친 정제 작업을 수행하였다. 최종 정제된 단어들은 텍스트 사이트에서 제공해 주는 키워드 빈도 분석을 통해 빈도가 높은 상위 70개의 주요 단어를 추출하였으며, TF-IDF와 연결 중심성 수치를 확인하였다. 여기서 TF-IDF는 TF(Term Frequency)와 IDF(Inverse Document Frequency)를 곱한 값으로 특정 단어가 한 문서 내에서 얼마나 중요한가를 나타내는 통계적 수치인데, 단순히 출현 빈도가 높다고 중요한 단어로 판단하기에 한계가 있을 때 확인하는 수치이다[2]. 마지막으로 연결 중심성 수치는 단어가 가지고 있는 연관성을 분석하는 데 활용하였다.

두 번째 단계는 의미 연결망 분석이다. 의미 연결망 분석은 키워드의 빈도와 연결 관계를 기반으로 네트워크 공간에 시각화하여 텍스트의 의미와 특성을 도출하는데 이용되는 방법이다[22]. 추출된 70개의 단어는 의미 연결망 분석을 위해 텍스트 사이트를 이용하여 1-mode 형태의 매트릭스 데이터로 변환하였다. 변환된 데이터는 Ucinet 6 프로그램을 이용하여 전체 네트워크에 대한 통계적 유의성 검정을 수행하였다. 마지막으로 Ucinet 6의 Netdraw 프로그램을 이용하여 전체 네트워크의 구조를 살펴보았으며, CONCOR 분석을 통해 텍스트들이 구성하고 있는 군집에 관해 확인하였다.

3. 연구 결과

3.1. 텍스트 마이닝 분석 결과

이 연구에서 ‘아두이노 강의’에 대해 텍스트 마이닝 분석을 수행한 결과 키워드의 빈도, TF-IDF, 연결 중심성은 <Table 1>과 같다. 빈도 분석 결과 아두이노(16,364), 강의(10,803), 활용(3,862), 교육(3,577), 코딩

(3,526), 학생(1,763), 수업(1,689), 기초(1,635), 진행(1,535), 스크래치(1,478) 등으로 나타났다. TF-IDF 분석 결과 교육(5,740.30), 코딩(5,657.12), 활용(5,538.16), 강의(4,305.69), 학생(3,923.09), 수업(3,863.29), 기초(3,745.94), 진행(3,456.50), 스크래치(3,375.10), 문의(3,228.87) 등으로 나타났다. 마지막으로 연결 중심성 분석 결과 아두이노(0.17859), 강의(0.14931), 활용(0.05839), 교육(0.05117), 코딩(0.04729), 학생(0.03754), 수업(0.03640), 진행(0.03044), 기초(0.02467), 시간(0.02230) 등을 확인할 수 있었다.

<Table 1> Result of text mining analysis

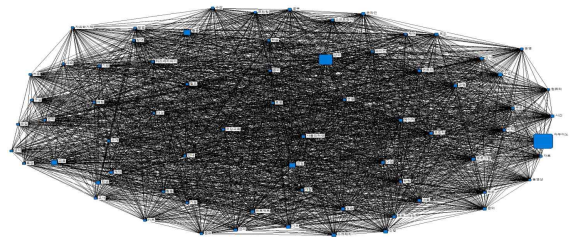
No	Keyword	frequency	Keyword	TF-IDF	Keyword	Degree centrality
1	Arduino	16,364	education	5,740.30	Arduino	0.17859
2	lecture	10,803	coding	5,657.12	lecture	0.14931
3	use	3,862	use	5,538.16	use	0.05839
4	education	3,577	lecture	4,305.69	education	0.05117
5	coding	3,526	student	3,923.09	coding	0.04729
6	student	1,763	class	3,863.29	student	0.03754
7	class	1,689	basic	3,745.94	class	0.03640
8	basic	1,635	progress	3,456.50	progress	0.03044
9	progress	1,535	scratch	3,375.10	basic	0.02467
10	scratch	1,478	inquiry	3,228.87	time	0.02230
11	program	1,105	program	2,927.46	program	0.02107
12	programming	988	programming	2,701.13	scratch	0.02088
13	inquiry	947	sensor	2,684.18	process	0.01994
14	time	923	time	2,628.07	instructor	0.01852
15	sensor	904	project	2,606.75	programming	0.01798
16	project	888	process	2,483.82	start	0.01719
17	process	863	printer	2,459.56	produce	0.01704
18	hands-on activity	847	maker	2,451.66	study	0.01704
19	printer	809	hands-on activity	2,444.72	project	0.01700
20	entry	804	entry	2,367.14	hands-on activity	0.01663
21	maker	803	Arduino	2,353.94	printer	0.01659
22	instructor	732	instructor	2,217.31	various	0.01637
23	study	722	study	2,213.78	contents	0.01480
24	online	721	produce	2,180.05	online	0.01480
25	produce	711	online	2,176.05	sensor	0.01445
26	various	687	various	2,052.26	teacher	0.01423
27	contents	645	board	2,042.07	ready	0.01394

28	board	639	Internet of Things	2,042.07	possible	0.01360
29	Internet of Things	636	contents	2,000.32	computer	0.01341
30	start	618	start	1,930.40	private institute	0.01312
31	videos	593	kit	1,911.05	software	0.01300
32	special lecture	587	printing	1,906.13	printing	0.01281
33	kit	582	videos	1,904.00	connect	0.01259
34	python	565	special lecture	1,891.46	development	0.01259
35	printing	551	python	1,873.99	robot	0.01246
36	connect	550	connect	1,827.73	introduce	0.01221
37	robot	534	robot	1,819.48	maker	0.01189
38	software	533	software	1,772.37	sign up	0.01183
39	operation	529	computer	1,740.32	board	0.01148
40	computer	520	operation	1,717.94	thinking	0.01139
41	development	491	development	1,671.83	python	0.01139
42	no charge	482	no charge	1,662.76	kit	0.01123
43	private institute	469	private institute	1,638.72	necessary	0.01117
44	ready	445	data	1,536.12	data	0.01076
45	possible	438	ready	1,528.74	no charge	0.01069
46	data	427	control	1,520.43	university	0.01066
47	control	420	possible	1,517.36	Internet of Things	0.01057
48	sign up	417	activity	1,505.87	coding education	0.01054
49	activity	414	sign up	1,473.08	guide	0.01044
50	teacher	412	teacher	1,471.93	method	0.01016
51	guide	407	recruitment	1,450.32	videos	0.01016
52	camp	403	guide	1,447.46	professor	0.01000
53	recruitment	400	camp	1,437.58	provide	0.00997
54	introduce	400	introduce	1,435.65	learning	0.00991
55	provide	396	provide	1,408.34	special lecture	0.00987
56	target	393	target	1,398.73	first	0.00987
57	learning	382	learning	1,377.45	introduce	0.00984
58	coding education	372	coding education	1,373.09	school	0.00984
59	base	367	experience	1,363.31	entry	0.00953
60	method	358	base	1,339.20	activity	0.00953
61	experience	358	university	1,322.19	target	0.00943
62	university	354	method	1,304.26	control	0.00934

63	necessary	344	necessary	1,268.69	present	0.00921
64	first	335	language	1,265.81	interest	0.00899
65	language	333	Raspberry H	1,261.26	operation	0.00864
66	Raspberry H	330	first	1,246.96	participation	0.00861
67	composition	320	composition	1,204.56	composition	0.00855
68	thinking	316	thinking	1,189.51	field	0.00849
69	participation	312	participation	1,185.99	grade level	0.00845
70	freemaster system	307	school	1,161.48	experience	0.00845

3.2. 의미 연결망 분석 결과

의미 연결망 분석을 위해 먼저 Ucinet 6 프로그램에서 전체 네트워크에 대한 통계적 유의성 검정을 수행하였다. 그 결과 z-score보다 절댓값이 큰 값이 나올 확률은 0.001로 나타났다. 이는 유의수준 5% 기준보다 작은 값이므로 네트워크 데이터 간에 통계적으로 유의한 관계가 있는 것을 알 수 있었다. 다음으로 Netdraw 프로그램을 이용하여 ‘아두이노 강의’ 관련 데이터를 시각화하여 나타내었는데 그 결과는 (Fig. 2)와 같다. 이 그림에서 확인할 수 있듯이 70개의 주요 키워드들은 복잡하게 연결된 형태로 파악되었으며, 대부분 키워드가 상호 연결된 것을 파악할 수 있었다.

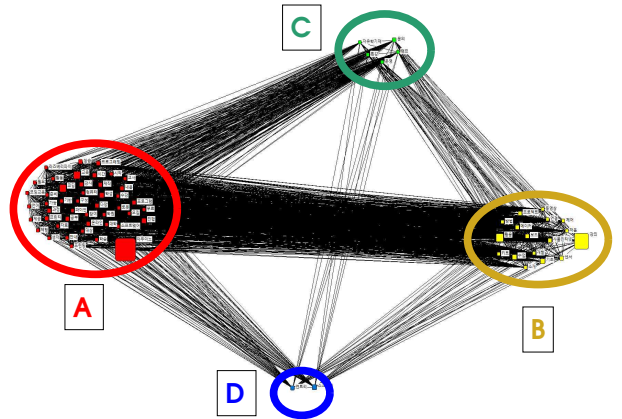


(Fig. 2) Result of network analysis

다음으로 아두이노 강의의 키워드를 중심으로 CONCOR(CONvergence of iterated CORrelation) 분석을 수행하여 이 키워드 간의 실제적인 유사 군집을 확인하였다. CONCOR 분석은 실제 단어들의 상관관계 분석을 반복적으로 수행하여 유사한 군집으로 분류함으로써

써 아두이노 강의와 관련된 키워드들의 좀 더 현실적인 요소들의 분류가 가능하다[11]. CONCOR 분석 결과 4 개의 군집이 형성되는 것으로 확인되었으며, 이 분석 결과는 (Fig. 3)과 같다.

첫 번째 군집은 아두이노, 수강, 실습, 공부, 대학, 코딩, 코딩 교육, 프로그래밍, 학생, 학습, 학원 등의 키워드로 구성되어 있었다. 이들은 아두이노와 관련된 다양한 교육 활동이었기 때문에 ‘아두이노 관련 교육’으로 명명하였다. 두 번째 군집은 강의, 수업, 메이커, 보드, 사물인터넷, 센서, 연결, 키트, 제어, 프로젝트 등의 키워드로 구성되어 있었다. 이들은 아두이노에서 피지컬 컴퓨팅 기능을 구현할 수 있는 키워드와 관련 있었기 때문에 ‘피지컬 컴퓨팅 관련 강의’로 명명하였다. 세 번째 군집은 문의, 운영, 캠프, 특강 등의 키워드로 구성되어 ‘아두이노 특강’으로 명명하였으며, 네 번째 군집은 스크래치, 엔트리 등의 키워드로 구성되어 ‘GUI(Graphical User Interface) 프로그래밍’으로 명명하였다.



(Fig. 3) Result of CONCOR analysis

<Table 2> Four clusters identified as a result of CONCOR analysis

Group	Clusters	Number	Keyword
A	Arduino-related education	46	activity, Arduino, base, coding, coding education, composition, computer, contents, data, education, experience, no charge, guide, hands-on, activity, instructor, language, learning, necessary, online, participation, possible, printer, printing, private institute, process, produce, program, programming, progress, provide, python, Raspberry Pi, ready, recruitment, robot, sign up, software, start, student, study, target, teacher, thinking, time, university, various
B	Physical computing-related lecture	17	basic, board, class, connect, control, development, first, Internet of Things, introduce, kit, lecture, maker, method, project, sensor, use, videos
C	Arduino special lecture	5	camp, free semester system, operation, inquiry, special lecture
D	GUI programming	2	entry, scratch

4. 논의

4.1. 텍스트 마이닝 분석 결과에 대한 논의

이 연구에서는 아두이노 강의에 대한 사회적 인식을 알아보고자 텍스트 마이닝 분석을 수행하였으며, 이 연구의 결과를 토대로 다음과 같이 논의하고자 한다.

첫째, ‘교육’이 빈도 4위(3,577), TF-IDF 1위(5,740.30), 연결 중심성 4위(0.05117)로 나타났다. 이 순위는 빈도나 연결 중심성에서 상위권이었으며, 특히 TF-IDF에서는 가장 높은 순위를 나타냈다. TF-IDF는 단어 등장 횟수의 단순 빈도를 통해 중요도를 측정하는 것이 아니라 비중 있게 다루어지는 단어를 찾는 방법으로 활용되므로[7], 이 키워드가 전체 키워드 중에서 가장 중요한 역할을 하는 키워드임을 확인할 수 있었다.

인터넷에서 텍스트 기록으로 남겨진 아두이노 강의와 관련된 ‘교육’ 키워드는 크게 두 가지 유형으로 생각해 볼 수 있다. 첫 번째 유형은 아두이노 자체를 강의한 교육 사례이다. 아두이노는 2005년 마사모 벤지 교수팀에 의해 개발된 마이크로컨트롤러 보드로, 그 종류가 다양하고 여러 가지 전자 기기를 연결하여 다양한 작업을 쉽고 편리하게 수행할 수 있어 큰 인기를 누리며 발전해 왔다[14]. 일반적으로 사용되고 있는 아두이노는 바로 이러한 보드를 의미한다. 그러나 아두이노는 코드뿐만 아니라 프로그래밍 언어, 프로그램을 사용하기 위한 개발 환경, 웹사이트(<http://www.arduino.cc>)나 워크숍

등을 포함한 전체를 일컫기도 한다[23]. 이처럼 아두이노는 다양한 의미로 사용되기 때문에 이와 관련된 강의 자료들이 '교육' 키워드와 함께 인터넷상에 공유되었던 것으로 유추할 수 있다.

두 번째 유형은 강의에서 아두이노를 다양한 교육 분야에서 활용한 사례이다. 아두이노는 가격이 저렴하며, 프로그램 작성을 위해 필요한 통합 개발 환경을 무료로 다운로드하여 사용할 수 있고, 무엇보다도 마이크로컨트롤러에 대한 배경지식이 없이도 창의적인 표현을 하고자 하는 학생이나 일반인에게 쉽게 적응하고 이용될 수 있도록 개발되었다[23]. 이에 중학교의 기술 교과, 특성화 고등학교의 전공 교과 등과 같이 특정 교과를 중심으로 한 교육 활동이나 STEAM 교육과 같이 여러 가지 교과에서의 융합을 주제로 한 교육 활동에서도 아두이노가 활용되고 있었다. 또한, 코딩이나 프로그래밍 교육, 메이커 교육, 영재교육, 발명 교육 등의 특정 주제를 기반으로 한 교육 활동이나 동아리 활동과 같은 교과 외의 활동에서도 활용되었다. 마지막으로 대학에서는 공학 교육 활동이나 캡스톤 디자인과 같이 종합적인 프로젝트 활동에서의 활용 사례[19]를 확인해 볼 수 있었고, 일반인을 대상으로는 창업 교육 활동[10]에서도 아두이노가 활용되고 있었다. 이처럼 아두이노는 다양한 교육 대상의 강의 활동에서 활용되고 있었기에 이와 관련된 수많은 기록이 인터넷상에 남아 교육이란 키워드가 빈도, TF-IDF, 연결 중심성 등에서 상위권의 순위를 나타낸 것으로 유추할 수 있다.

둘째, '코딩'이 빈도 5위(3,526), TF-IDF 2위(5,657.12), 연결 중심성 5위(0.04729)로 나타났다. 코딩 역시 TF-IDF에서 검색 키워드보다 높은 순위를 나타낸 키워드로 인터넷상 기록으로 남겨진 '아두이노 강의' 관련 텍스트와 매우 밀접한 관련이 있는 키워드로 볼 수 있다. 코딩은 인간의 언어를 컴퓨터가 이해할 수 있는 언어인 각종 코드(code)를 이용하여 나타내는 프로그램을 작성하는 것으로[3], 선진국에서는 벌써 수년 전부터 코딩의 중요성이 강조하여 왔다. 예를 들어, 미국의 오바마 전 대통령은 '컴퓨터 과학 교육 주(week)'의 축사에서 게임을 즐기기도는 직접 해당 프로그램을 만들어 보라고 언급하였고, 애플 창업자인 스티브 잡스도 미래 사회를 준비하려면 코딩을 배워야 함을 여러 차례 강조하였다[3]. 이러한 배경에 따라 미국에서는 2015년

워싱턴 DC와 시카고 및 17개 주에서 코딩을 필수 교육으로 지정한 바 있다[25]. 그 외 영국, 중국, 인도, 일본, 핀란드 등의 나라에서도 코딩의 중요성을 인식하여 이를 필수 이수 과목으로 지정하는 등 국가적 차원에서 코딩 교육 관련 교육 정책을 추진하고 있다[20, 24-25]. 우리나라에서도 4차 산업혁명에 맞춘 교육과정의 개편 요구와 SW 교육의 확대라는 세계적인 추세에 맞추기 위해 교육과정에서 코딩을 도입하였다[25]. 이처럼 코딩에 대한 중요성이 강조되고 있는 시대적 흐름에 따라 다양한 코딩 관련 강의가 이루어졌는데, 아두이노는 이러한 코딩을 체험할 수 있는 좋은 교구였다. 이는 아두이노 사이트에서 무료로 다운로드 할 수 있는 통합 개발 환경 프로그램인 스케치 프로그램이 C언어를 기반으로 하고 있으며[5], 그 외에 파이썬, 스크래치, 엔트리와 같은 코딩을 작성할 수 있는 외부 프로그램과도 연동이 가능하기 때문이다. 이러한 배경에서 아두이노 강의 관련 인터넷 기록에서 코딩 관련 키워드가 자주 등장했기 때문에, 이 키워드의 빈도나 TF-IDF, 연결 중심성이 높았던 것으로 해석할 수 있다.

셋째, '문의'가 빈도 13위(947), TF-IDF 10위(3,228.87), 연결 중심성 117위(0.00596)로 나타났다. 이 키워드는 빈도나 TF-IDF보다 연결 중심성이 매우 낮은 키워드였다. 여기에서 연결 중심성이 매우 낮다는 것은 다른 키워드들과 노드 간 연결이 되지 않는 독립된 키워드라는 것을 의미한다. 이 키워드를 포함하고 있는 인터넷 텍스트의 원문을 확인해 본 결과, 이들은 컴퓨터나 코딩 교육 학원, 각종 교육 기관, 아두이노 관련 실습 재료 판매처 등에서 작성한 광고성 텍스트임을 알 수 있었다. 빅데이터 분석 방법으로 사회적 인식을 조사한 선행 연구의 경우 이와 같은 광고성 키워드를 삭제하여 분석한 사례도 있었다. 그러나 본 연구에서는 이러한 광고 관련 텍스트 역시 사회적 인식을 반영하고 있는 것으로 보고 이를 포함하여 분석하였으며, 이를 통해 의미 있는 시사점을 확인할 수 있었다. 예를 들어, '문의' 키워드가 포함된 텍스트는 컴퓨터나 코딩 학원에서 작성된 사례가 많았는데, 그들은 자신의 학원에서 개설된 아두이노 관련 강의 정보를 제시하고 있었으며 이를 통해 수강생을 모집하려는 의도가 있었다. 이는 사교육 시장에서 아두이노 강의를 개설한 것을 홍보한 광고 글이 많이 등장했음을 유추할 수 있는 결과이다. 2015 개정

교육과정에서 소프트웨어 교육이 필수 과목으로 지정될 무렵 이와 관련된 사교육을 부추길 것이라는 우려가 있었는데, 인터넷상의 기록을 통해 이러한 우려가 현실로 나타났음을 확인할 수 있었다.

넷째, ‘센서’가 빈도 15위(904), TF-IDF 13위(2,684.18), 연결 중심성 22위(0.01445)로 나타났다. 이 키워드 역시 연결 중심성보다 빈도, TF-IDF의 순위가 상대적으로 높았는데, 이는 센서란 키워드가 다른 키워드들과 연관성은 상대적으로 낮았지만 개별적으로는 중요한 의미를 가지는 것을 의미한다. 센서는 로봇, 자동차, 비행기 등 대부분의 시스템을 효율적으로 동작하도록 제어하여 자동화를 가능하게 하고, 각 부품에 대한 기능의 효율을 높이며, 사용자에게 편리성과 생활의 유용성을 제공하기 위한 용도로 활용되는 중요한 부품이다[14]. 자주 활용되는 센서로는 조도 센서, 적외선 센서, 초음파 센서, 기울기 센서, 진동 센서, 온도 센서, 가속도 센서, 자이로스코프 센서 등이 있다. 이들 센서를 이용하면, 다양한 프로젝트를 진행할 수 있다. 예를 들어, 기울기 센서를 이용하면 자동차나 로봇의 움직임을 제어할 수 있으며, 적외선 센서를 이용하면 센서로 사람이나 동물의 움직임을 감지하여 전등을 켜거나 소리가 나게 할 수 있다[6]. 아두이노를 이용한 대부분의 프로젝트에서는 이들 각종 센서로부터 데이터를 수집 받아 정보를 수집하여 이를 처리하는 과정이 포함하고 있다. 따라서 인터넷상에서는 아두이노 강의에서 센서를 다루는 방법과 관련된 많은 기록물이 쌓였던 것으로 보이며, 센서란 키워드는 상대적으로 다른 키워드들과 적게 연계되어 기록되었음을 확인할 수 있었다.

4.2. 의미 연결망 분석 결과에 관한 논의

이 연구에서는 의미 연결망 분석을 위해 네트워크 시각화 및 CONCOR 분석을 수행하였으며, 그 결과 4개의 군집을 확인할 수 있었다. 각 군집은 ‘아두이노 관련 교육’, ‘피지컬 컴퓨팅 관련 강의’, ‘아두이노 특강’, ‘GUI 프로그래밍’ 등으로 명명하였다.

첫째, ‘아두이노 관련 교육’ 군집의 주요 키워드는 아두이노, 수강, 실습, 공부, 대학, 제공, 제작, 체험, 코딩, 코딩 교육, 프로그램, 프로그래밍, 학생, 학습, 학원 등이었다. 이 군집에 해당하는 키워드는 총 46개로 4개의

군집 중 가장 많은 키워드를 포함하고 있었으며, 이들 키워드는 대부분 아두이노 교육과 관련되어 있었다. 예를 들어, 학생, 대학은 아두이노 교육의 대상자를 의미하고 있었고, 공부, 수강, 실습, 제작, 체험, 학습은 아두이노의 교육과정에서 학습자가 행하는 활동과 관련이 있었다. 앞 절에서 확인한 바와 같이 ‘교육’ 키워드가 빈도, TF-IDF, 연결 중심성 등에서 상위권 순위를 나타내어 이 키워드가 ‘아두이노 강의’를 표현하는 핵심적인 키워드임을 확인하였는데, 군집 분석을 통해서도 ‘아두이노 강의’와 관련된 키워드들은 ‘아두이노 교육’이란 주제로 하나의 군집을 형성하고 있음을 확인할 수 있었다. 이러한 연구 결과를 통해 인터넷에 기록으로 남아 있는 ‘아두이노 강의’ 관련 키워드들은 아두이노 교육과 관련된 주제가 다루어지고 있음을 확인할 수 있었다. 실제 발표된 학술 연구에서도 아두이노를 교육 분야에 활용한 다양한 연구가 보고되고 있으며, 아두이노 적용 교육의 효과를 측정하는 실험 연구도 다수의 연구자에 의해 수행되었다. 특히, 국내에서 수행된 아두이노 적용 교육의 효과 크기를 확인한 메타 분석 연구도 수행되었는데, Lee[17]의 연구에서는 0.656의 효과 크기를, Jang[8]의 연구에서는 0.537의 효과 크기를 확인하였다. 일반적으로 메타 분석 연구는 해당 분야의 실험 연구가 어느 정도 축적되었을 때 수행되므로 이들 연구를 통해 국내에서 아두이노 교육에 대한 연구가 어느 정도 축적되었음과 아두이노를 교육에 적용하였을 때 긍정적 효과가 있었음을 확인할 수 있었다. 이를 볼 때 아두이노 교육에 대한 학문적 기초가 형성되어 많은 연구자가 이에 대한 학술 연구를 진행하는 것과 유사하게, 인터넷상에서도 일반 대중들이 아두이노 교육과 관련된 여러 가지 활동들에 대한 기록들을 남겼음을 확인할 수 있었다.

둘째, ‘피지컬 컴퓨팅 관련 강의’ 군집의 주요 키워드는 강의, 수업, 메이커, 보드, 사물인터넷, 센서, 연결, 키트, 프로젝트, 제어 등 이었다. 여기서 피지컬 컴퓨팅이란 프로그램이나 센서 등을 이용하여 컴퓨터가 인간의 감각 기관 역할을 하거나 그에 반응하도록 하는 것으로, 키보드나 마우스 같은 입력 방법 대신 소리, 동작, 빛, 열 등의 물리적인 신호를 감지하는 방법으로 정보를 입력하고 표현하는 방법을 말한다[12]. 아두이노는 내장된 GPIO(general-purpose input/output) 장치를 이용하여 다양한 센서 및 액추에이터를 제어할 수 있는데, 이를

구성하기 위해서는 아두이노 키트나 보드가 필요하며, 이에 대한 회로 구성 방법을 익혀야 한다. 이에 아두이노 강의 중 이러한 피지컬 컴퓨팅과 관련된 주제를 다룬 내용이 인터넷상에 많이 탑재되어 이들이 하나의 주제를 형성하고 있음을 확인할 수 있었다.

셋째, ‘아두이노 특강’ 군집은 문의, 운영, 캠프, 특강, 자유학기제 등 5개의 키워드로 구성되었고, ‘GUI 프로그래밍’은 스크래치, 엔트리 등 2개의 키워드로 구성되었다. 이들은 많은 키워드를 포함하지는 않았지만, 이들 키워드가 의미하는 바를 분명하게 파악할 수 있었다. 즉, 아두이노 특강 군집에 해당하는 키워드들로는 아두이노 관련 강의가 특강이나 캠프로 운영되는데, 이에 대한 정보가 필요하다면 문의하라는 내용을 유추할 수 있었다. 이를 통해서 아두이노와 관련된 특강이나 캠프를 홍보하는 내용이 하나의 주제로 다루어 지고 있음을 확인할 수 있었다. GUI 프로그래밍 군집에 해당하는 스크래치와 엔트리는 아두이노 프로그래밍을 작성할 수 있는 대표적인 그래픽 기반 프로그래밍 도구이다. 특히 이러한 그래픽 기반 프로그래밍 도구는 텍스트 기반 프로그래밍 도구보다 프로그램 작성 시 오류 가능성이 적기 때문에, 초급자들을 위한 아두이노 교육 활동에서 많이 활용되고 있었다. 이처럼 초급자들을 위한 프로그래밍 언어를 소개한 인터넷 텍스트 내용이 서로 연계가 되는 부분이 많았기 때문에, 이들 키워드가 하나의 군집으로 형성되었음을 확인할 수 있었다.

5. 결론

이 연구는 인터넷 기반의 텍스트 자료를 활용한 빅데이터 분석을 통해 대중의 아두이노 강의에 관한 인식을 확인하였다. 이를 위해 네이버 사이트의 블로그, 카페, 뉴스 채널의 텍스트 자료를 ‘아두이노 강의’ 키워드로 수집하여 텍스트 마이닝과 의미 연결망 분석을 수행하였다.

텍스트 마이닝 분석은 텍스트 사이트를 활용하였으며, 빈도 분석 결과 아두이노, 강의, 활용, 교육, 코딩, 학생, 수업, 기초, 진행, 스크래치 등으로 순으로 나타났고, TF-IDF 분석 결과 교육, 코딩, 활용, 강의, 학생, 수업, 기초, 진행, 스크래치, 문의 순으로 나타났다. 연결 중심성 분석 결과 아두이노, 강의, 활용, 교육, 코딩, 학생, 수

업, 진행, 기초, 시간 등의 순으로 나타났다. 연구 결과 교육과 코딩 키워드는 TF-IDF 분석 순위에서 최상위권에 속한 중요 키워드임을 확인하였고, 문의, 센서 등의 키워드들은 연결 중심성 분석 순위보다 TF-IDF 분석 순위가 높은 키워드임을 확인하였다.

의미 연결망 분석은 Ucinet 6과 Netdraw를 활용하여 네트워크의 시각화 및 CONCOR 분석을 수행하였다. CONCOR 분석 결과 4개의 군집을 확인할 수 있었다. 첫 번째 군집은 아두이노, 수강, 실습, 공부, 대학, 코딩, 코딩 교육 등의 키워드로 구성된 ‘아두이노 관련 교육’ 군집이었고, 두 번째 군집은 강의, 수업, 메이커, 보드, 사물인터넷, 센서, 연결, 키트 등의 키워드로 구성된 ‘피지컬 컴퓨팅 관련 강의’ 군집이었다. 세 번째 군집은 문의, 운영, 캠프, 특강 등의 키워드로 구성된 ‘아두이노 특강’ 군집이었고, 네 번째 군집은 스크래치, 엔트리 등의 키워드로 구성된 ‘GUI(Graphical User Interface) 프로그래밍’ 군집이었다.

이 연구를 통해 인터넷상에서 아두이노 강의와 관련하여 일반 대중들의 여러 가지 의미 있는 사회적 인식을 확인할 수 있었다. 특히, 일부 키워드의 경우는 사교육 업체에서 자신들의 강의 정보를 홍보하기 위해 작성된 것으로 유추할 수 있었는데, 이러한 분석 결과는 빅데이터 분석을 통해서만 확인할 수 있는 정보였다. 인터넷은 누구나 자유롭게 글을 남길 수 있는 공간이기 때문에 아두이노 강의에 대한 일반적인 정보뿐만 아니라 아두이노 강의를 개설한 각종 학원의 홍보성 글도 남게 된다. 2015 개정 교육 과정부터 시행된 소프트웨어 교육의 의무화 정책 추진 당시 이와 관련된 사교육 시장이 과열될 것이라는 우려가 있었는데, 이 연구의 빅데이터 분석을 통해 사교육 시장에서 실제로 반응한 사례를 확인할 수 있었다. 이 외에도 이 연구에서 확인한 여러 가지 아두이노 강의와 관련된 키워드들을 정밀하게 분석해 본다면 소프트웨어 교육과 같이 최근 정부에서 중점적으로 추진하고 있는 교육 정책들이 사회에서 어떻게 반응하고 있는지를 파악할 수 있을 것이다. 이 연구의 결과는 아두이노 강의를 준비하는 교수자나 해당 주제를 연구하는 연구자, 나아가 소프트웨어 교육이나 코딩 교육과 관련 정책을 수립하는 정책 입안자들에게 의미 있는 시사점을 제공하는 자료로 활용될 것이다.

참고문헌

- [1] Choi, H. & Choi, Y.(2016). A study on children's creativity and character based on big data. *Journal of Childrens Literature and Education*, 17(4), 601-627.
- [2] Choi, J. & Park, S.(2020). A study on perception of golf lesson using big data analysis. *Journal of Golf Studies*, 14(1), 151-163.
- [3] Choi, S.(2019). Investigating social perception of 'young children's coding' using big data. *The Journal of Humanities and Social science*, 10(6), 1589-1603.
- [4] Gu, Y.(2020). Trend analysis on clothing care system of consumer from big data. *Fashion & Textile Research Journal*, 22(5), 639-649.
- [5] Hur, K.(2018). An educational application of an IoT device development task for visualization of computational thinking. *The Journal of Education*, 38(3), 255-267.
- [6] Hur, K., Lee, J., Lee, H. & Lee, H.(2015). Development of education program for physical computing using Arduino n-screen communication boards *Journal of Practical Engineering Education*, 7(2), 97-105.
- [7] Hwang, Y.(2019). A study on the recognition of universal design through text mining. *Journal of the Korea Institute of Spatial Design*, 14(1), 57-66.
- [8] Jang, B.(2021). The meta-analysis on effects of Arduino-based education for secondary school students. *Journal of Industrial Convergence*, 19(3), 61-65.
- [9] Kim, D., Park, S., Park, K. & Yang, K.(2018). Digitization of color of solution using Arduino. *Journal of Science Education for the Gifted*, 10(3), 244-255.
- [10] Kim, K., Ahn, G., Lim, H. & Jung, D.(2015). Research about senior citizen IT start-up education linking the IoT. *Journal of the Korea Institute of Information and Communication Engineering*, 19(11), 2710-2716.
- [11] Kim, K. & Kim, E.(2018). A study on social perception of forest education for young children through social network analysis based on big data. *The Journal of Korea Open Association for Early Childhood Education*, 38(5), 287-308.
- [12] Kim, Y. & Hong, K.(2016). The effects of physical computing based software applications using Arduino on logical thinking of elementary school students. *Korean Journal of Thinking Development*, 12(2), 47-72.
- [13] Kim, Y., Kim, T. & Kim, J.(2015). Development and application of programming education program of robot for improvement of elementary school girls' creativity. *Journal of The Korean Association of Information Education*, 19(1), 31-44.
- [14] Koh, B.(2016). A study on the STEAM education based Arduino. *The Journal of Education Studies*, 53(4), 1-18.
- [15] Koh, J. & Chong, Y.(2020). Analysis of Taiwanese food trends in Korea using social big data: Focusing on blog text analysis. *Journal of Tourism Management Research*, 96(-), 71-91.
- [16] Kwon, J. & Park, J.(2020). Elementary safety education awareness through big data. *Journal of Learner-Centered Curriculum and Instruction*, 20(16), 1351-1371.
- [17] Lee, E.(2020). A meta-analysis of the effects of arduino-based education in Korean primary and secondary schools in engineering education. *European Journal of Educational Research*, 9(4), 1503-1512.
- [18] Lee, J. & Lee, B.(2017). A study on the promotion of competitiveness of technology education based on the 4th industrial revolution in Gyeong-nam area : Practical using Arduino convergence contents education program. *The Journal of Northeast Asia Research*, 32(2), 279-299.
- [19] Lee, J. & Lee, S.(2018). Development of creative

convergence education program for engineering college students. *The Korean Society of Computer and Information*, 23(5), 15-23.

- [20] Lee, K.W. & Nam, H.W.(2018). The effect of coding education on elementary school students' computational thinking. *Journal of Korean Practical Arts Education*, 31(4), 269-284.
- [21] Noh, Y., Chang, H. & Park, B.(2020). Types and trends of differentiated cafes by sematic network analysis methods. *Journal of the Table & Food Coordinate*, 15(3), 41-58.
- [22] Park, K.(2019). Perception on 'Sungsu-Dong' through big data analysis. *Journal of Cultural Product & Design*, 58(-), 45-52.
- [23] Son, K.H. & Sohn, W.S.(2014). The development and application to computer programming education using Arduino. *The Journal of Education*, 34(3), 159-179.
- [24] Song, Y.(2017). Design and implementation of reflection-based coding education: Case study of "SW and computational thinking" courses at H university. *Journal of Educational Technology*, 33(3), 709-736.
- [25] Suh, E.(2017). Development of creative thinking and coding course method on design thinking using flipped learning. *Journal of Learner-Centered Curriculum and Instruction*, 17(16), 173-199.
- [26] Yoon, S. & Jang, E.(2014). The application of micro controller board to engineering education for multidisciplinary capstone design. *Journal of Digital Convergence*, 12(2), 531-537.

저자소개



이 은 상

2003 한국교원대학교 기술교육과 (학사)

2013 충남대학교 기술교육과 기술교육전공(석사)

2015 충남대학교 기술교육과 기술교육전공(박사)

관심분야: AI교육, 빅데이터, 저비용 마이크로컨트롤러, 발명교육

e-mail: eslee@kongju.ac.kr