

<http://dx.doi.org/10.17703/JCCT.2023.9.2.381>

JCCT 2023-3-48

빅데이터 분석을 활용한 지역 맞춤형 교육프로그램 선정 모형 개발

A Study on Regional-customized education program selection model using big data analysis

김현성*, 김진숙**

Hyeon-Seong Kim*, Jin-Sook Kim**

요약 본 연구는 빅데이터 분석을 활용한 지역 맞춤형 교육프로그램 선정 모형 개발을 주요 목적으로 한다. 우선, 문헌 고찰을 통해 빅데이터 및 교육의 개념 및 특성 그리고 빅데이터 기술과 연구 활용 등의 이론을 분석하여, 이를 평생교육 빅데이터 활용을 위한 선결과제와 기초 연구자료로 제공한다. 아울러 교육 데이터 수집의 방법과 교육의 특성에 적절한 빅데이터 활용 방법을 제시하고 이를 활용한 지역 맞춤형 교육프로그램 선정 모형을 개발하였다. 지역 맞춤형 교육프로그램 선정 모형 개발은 총 6단계로 진행되었다. 본 연구에서 제시한 맞춤형 교육프로그램 모델은 실질적 활용 면에 있어, 국가승인통계인 ‘평생학습 개인 실태조사’ 처럼 1년 후에 분석하지 않고 실시간으로 데이터가 제공되는 방식으로 활용 부분에 있어서도 선택적 분석이나 미래예측 등 자유도가 매우 높아 교육 분야에 빅데이터가 충분한 필요성과 가치가 있음을 알 수 있다. 뿐만 아니라 표본 모형에 사용되고 있는 모든 프로그램은 무료로 제공되고 있으며, 프로그래밍 특성상 커뮤니티 또한 활발하게 교류가 이루어지고 있어 추후 수정 및 보완 시에도 매우 용이하여 더욱 완성도 높은 교육프로그램 개발 모형을 개발할 수 있다.

주요어 : 빅데이터분석, 프로그램개발, 교육프로그램, 지역 맞춤형

Abstract This thesis is purposed to develop a regional-customized education program selection model using big data analysis. Based on the literature review, the concepts and characteristics of big data and lifelong education are analyzed. In addition, this thesis presents how to collect the data for lifelong education and to use big data suitable for the characteristics of lifelong education. Based on these results, a regional-customized lifelong education program selection model is developed. The regional customized lifelong education program model is developed by the following six steps. The customized education program model proposed in this study has a high degree of flexibility in terms of practical use, as it can be utilized in real-time data provision methods such as the nationally approved Lifelong Learning Personal Status Survey without the need for analysis one year later, allowing for selective analysis and future predictions. It is clear that there is a significant need and value for big data in the education field. Furthermore, all programs used in the sample model are provided free of charge, and due to the programming nature, the community is actively engaged in exchanges, making it very easy to modify and improve for the development of a more complete education program model in the future.

Key words : Education Program, Big Data Analysis, Program Development, Regional customized

정회원, 경북대학교 IT대학 교육연구부장
정회원, 대구한의대학교 평생교육융합학부 교수(교신저자)
접수일: 2023년 2월 15일, 수정완료일: 2023년 3월 3일
게재확정일: 2023년 3월 13일

Received: February 15, 2023 / Revised: March 3, 2023

Accepted: March 13, 2023

**Corresponding Author: jskim@dhu.ac.kr

Dept. of Secondary Special Education, Daegu Haany Univ,
Korea

I. 서론

현재 평생교육 기관의 평생교육사들은 프로그램 개발 방법으로 주로 전문 강사들에게 프로그램 커리큘럼을 제공받는 것을 채택한다. 지역의 평생교육 전문가와 기관 내 담당자가 커리큘럼을 만든 강사는 면접을 통해 평가하고 커리큘럼이 결정되면 강사는 홍보를 통해 학습자를 모집한다. 만약 정원을 채우지 못한다면 폐강하게 되는데, 이 경우 계획된 시간과 시설의 활용도는 떨어진다. 과거 평생교육프로그램은 다양한 관점을 기준으로 하여 개발되었으며, 그 중 가장 합리적 과정으로 인정받은 Yang과 Cervero(2001)의 전통적 관점의 프로그램 개발 과정은 보통 훈련이 얼마나 또 어떻게 필요한지 인식하는 것에서부터 학습 목표를 설정하고, 그 목표 달성을 위해 학습경험의 선정과 조직화를 거쳐서 프로그램을 실행 후 평가를 한다. 그러나 전통적 관점에서의 프로그램개발과정은 프로그램 개발자에 의해서 수행되는데 평생교육 기관의 특성, 이념, 행위자의 가치관, 사회적 상황 등에 주목하지 않고 현실에 맞지 않는 원리 혹은 절차에 주목하고 있다. 실제 프로그램 개발은 규범적으로 제시한 선형적 접근을 너무 강조한 나머지 전통적 관점에서 생각하는 선형적 절차와 순서로 이루어지지 아니하고 현실적 접근에 소홀하다는 비난이 있다 [4][9][10]. 그 외에도 상황 관점, 정치협상 관점, 비판적 관점 등 다양한 개발 이론이 있었고, Warren(2003)은 Houle, Kowalski, Caffarella, Tyler 등 대표적인 교육학자들의 프로그램 개발 모델을 분석해 본 결과, 프로그램 개발에 받아들여질 만한 간단명료한 정의와 단일화되고 정형화된 모델 그리고 전체를 포괄하는 모델을 찾지 못하였다 [17]. 이와 관련하여 Caffarella(2002)는 interactive model을 제시했고, 이 모델은 시작과 끝이 없는 형태로 Cervero와 Wilson의 투쟁(conflicting), 윤리적 의사결정, 전통적 관점에서의 원리를 담고 있다 [2].

김진화(2003)는 평생교육프로그램 개발의 내용·절차·의사결정·교육내용·정치·설계적 측면을 통합하여 협상 전문가, 내용 전문가, 의사결정 전문가, 과정 전문가, 학습설계 전문가로 역할을 수행하는 통합적인 관점을 제시했다. 이처럼 교육프로그램 개발은 지역사회와 함께 성장하고, 주민들이 주도적 방식으로 선정하여야 한다. 즉, 평생교육 문화를 정착시키기 위해 지역사회 주민들

의 욕구가 충분히 반영되면서 함께 성장할 수 있도록 평생교육프로그램 선정 및 개발하여 빅데이터를 활용한 개발 모형을 만들어야 한다. 하지만, 현재 빅데이터 연구는 공공, 금융, 기업 경영 분야, 통신·미디어 등 다양한 분야에서 인정받는 것에 비해 평생교육 분야에서의 연구 및 활용은 미비한 실정이다 [9].

따라서 본 연구는 문헌 고찰의 방법으로 교육프로그램 빅데이터 활용을 위한 선결과제와 그 연구의 기초자료로 제공하고, 교육프로그램 데이터 수집 방법과 교육 특성에 적합한 빅데이터 활용 지역 맞춤형 교육프로그램 선정 모형을 개발하고자 한다.

II. 연구 방법

1. 연구모형

연구모형은 연구 단계별에 따른 절차를 기반으로 문헌고찰, 분석레이어, 평가로 구분하며 다음 그림 1과 같다.

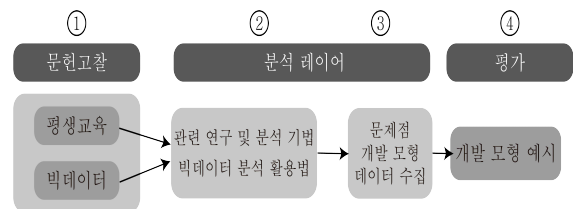


그림 1. 연구모형
Figure 1. research model

문헌고찰을 통한 빅데이터 분석 문헌조사(개념, 활용, 특성, 연구, 사례, 기술, 기획, 정책, 문제 및 한계점 등)와 교육 문헌조사(개념, 분류 체계, 특성, 정책, 문제 및 한계점 등)를 검토, 정리하여 대략적인 빅데이터 활용 지역의 맞춤형 교육프로그램 모형을 구축하였다. 이후 분석 레이어는 API(Application Programming Interface), python(모형 표본 프로그래밍 언어), pytrends(Google Trends), VBA(Visual Basic for Applications), 시각화 시트(빅데이터 센터 LOD클라우드)를 분석 도구로 선정하고, 자료를 분석하여 교육 데이터 수집의 방법, 교육 빅데이터 분석의 활용, 빅데이터를 활용한 지역 맞춤형 교육프로그램 모형을 확정하였다. 평가는 마지막으로 개발 모형에 대한 예시와 함께 평가를 진행하였다 [1][7][10].

2. 연구절차

구분 연구모형에 따른 연구절차는 문헌고찰, 분석도구 선정, 자료분석, 평가로 총 4단계로 표 1 과 같이 진행하였다.

표 1. 연구 단계별 절차
 Table 1. Research process steps

연구단계	세부내용
1단계	· 교육 빅데이터 생성과 수집방법 분석 · 빅데이터 문헌 분석
2단계	· 분석과 관련된 교육 연구와 빅데이터 분석 활용 방법
3단계	· 지역 맞춤형 교육프로그램 선정 모형 개발과 데이터 수집의 과정 제시 · 지역 교육프로그램 문제점
4단계	· 선정 모형 개발 예시와 평가

3. 교육 데이터 생성 및 수집방법

1) 교육 빅데이터 생성

교육으로 생성된 데이터 자료는 학습 영역별로 나눌 수 있으며, 영역별 교육데이터 생성 내용은 다음 표 2 와 같다.

표 2. 연구 단계별 절차
 Table 2. Content of generating data by learning area

구분	주요 내용
형식교육	초등학교, 중·고등학교, 대학(교), 사이버 원격대학, 방송통신대학교, 대학원(석·박사)에서 생성되고 있는 교육 데이터: 교육의 종류, 교육만족도, 수강인원 등 정규교육과정 전반과 관련된 데이터가 생성
비형식교육	대학 교육원, 주민자치센터, 평생학습관, 백화점 문화센터, 복지회관 등 교육기관에서 생성된 교육 데이터로 교육과 관련한 가장 많은 빈도의 데이터가 생성
무형식학습	별도의 교수자가 없는 자기주도적 학습으로, 인터넷 활동을 통한 데이터 생성을 제외하고는 수집에 어려움이 있음

교육부, 한국교육개발원(2018)에서 조사한 『2018 한국 성인의 평생학습실태』에 따르면, 한국 성인 교육 참여율은 42.8%이고 참여하는 인원중 비형식교육이 41.8%, 형식교육이 2.7%이며 교육 참여로 생성된 데이터는 학습 영역 가운데 비형식교육에 해당하는 교육기관으로부터 제한적으로 생성된다고 본다. 이외 교육에 활용할 수 있는 여러 데이터를 명확화할 수 없지만 인터넷에 생성되는 방대한 빅데이터를 자동수집방법인 크롤링(crawling), RSS Reader Open API 기법, 로그 수집기, 센싱(sensing) 기법을 활용하여 분산된 다양한 데이터를 수집하는 과정이 한정적 교육 데이터 생성 문제를 개선하고, 이는 교육 빅데이터 연구를 위해 필요하다.

2) 교육 빅데이터 수집방법

교육 데이터의 수집방법은 제한적으로 생성되는 교육기관들로부터 축적된 정형 데이터를 제공받아 수집하는 방법과 크롤링(crawling)기법 가운데 웹 로봇(Web Robot)을 활용해서 각 기관의 웹사이트 내 모든 페이지를 확인하고 이를 색인해 정형·비정형 데이터를 모두 포함한 전체 데이터를 수집하는 방법도 있다.

4. 교육 빅데이터 분석 활용

1) 기획단계

교육 데이터 생성 및 수집을 완료한 시점에 교육 기획 단계에서 교육기관 데이터를 활용할 경우 상·하향식 접근방식이 사용될 것으로 보이나 이외 생성되는 모든 교육 데이터의 원천을 파악하기에 어려움이 있을 것으로 생각되고, 기존의 연구가 미비한 실정에서 연구 결과 확인과 개선을 반복적으로 하여야 하는 부분을 볼 때, 교육의 데이터는 주로 프로토타이핑 접근방식을 사용해야 할 것으로 보여진다.

2) 데이터 마이닝 활용

교육기관 데이터를 활용할 경우 정형 데이터 분석을 활용하여 의사결정 나무 기법, 연관 관계분석 기법, 사례 기반추론, 인공지능망 기법, 군집분석 기법 등을 사용하고 그 외에 수집된 데이터는 비정형 데이터 분석을 활용한 텍스트마이닝, 웹마이닝, 오피니언마이닝, 데이터 시각화, 소셜 네트워크 분석 기법 등을 사용한다.

3) 교육 영역별 분석 유형

빅데이터를 활용한 교육 영역별 분석 유형의 예시는 6 가지로 다음 표 3 과 같다.

표 3. 교육 영역별 분석 유형
 Table 3. Types of analysis by educational area

분야	의사결정 유형(예시)	분야	의사결정 유형(예시)
프로그램 개발	교육프로그램 요구과약 학습자 선호 교육의 형태 파악 특정프로그램 효과를 분석	데이터 트래픽	교육의 만족도 평가 교육의 만족도 효율성 측정 다양한 재무를 예측
데이터 유형	교육 수준의 관리 교육 효과를 예측 교육의 문제점 및 커리큘럼 조정 교육장과 교보재 관리	교육 수요 관리	교육 수강생 결정 일자리 창출 목표 교육정보의 습득 경로 교육프로그램의 홍보
인적 자원 관리	이직 직원을 예측 교수자 채용 시기를 결정 직원보상액 산정 이직원인 분석과 대응 직원 교육 효과를 분석	마케팅 관리	교육비를 결정 교육 목표를 설정 교육기관 지점 위치 선정 온라인 미디어 광고 채널 선정 맞춤형 웹사이트를 구축

III. 연구결과

1. 지역 맞춤형 교육프로그램 선정 모형 개발

빅데이터를 활용한 교육프로그램 선정 모형 개발 과정은 다음 그림 2 와 같다.

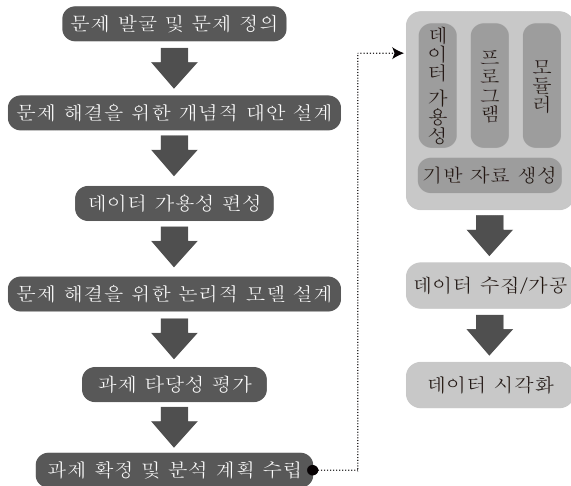


그림 2. 빅데이터를 활용한 교육프로그램 선정 모형 개발 과정
Figure 2. Development process of a model for selecting educational programs using big data

1) 문제 발굴 및 문제 정의

교육 기관의 교육사들은 프로그램 개발 방법으로 주로 전문 강사들에게 프로그램 커리큘럼을 제공받는 것을 채택한다. 지역의 교육 전문가와 기관 내 담당자가 커리큘럼을 만든 강사는 면접을 통해 평가하고 커리큘럼이 결정되면 강사는 홍보를 통해 학습자를 모집한다. 만약 정원을 채우지 못한다면 폐강하게 되는데, 이 경우 계획된 시간과 시설의 활용도는 떨어지는 문제

2) 문제 해결을 위한 개념적 대안 설계

교육이 지역사회와 함께 성장하는 주민 주도적 방식의 프로그램을 선정하여야 한다. 즉 교육 문화를 정착시키기 위해 교육프로그램 개발 및 선정의 과정은 지역사회 주민의 욕구를 충분히 반영하고 함께 성장 가능한 개발 모형이 필요하다. 따라서 교육 전문 강사들로부터 커리큘럼을 제공받는 방법이 아닌 주민의 의견이 충분히 반영된 모형으로 빅데이터를 활용한 교육프로그램 선정 모형 개발 모형을 제시

3) 데이터 가용성 편성

Naver DataLap과 Google Trends에서 크롤링(crawling)기법을 활용한 데이터를 수집할 것이며, 내용은 '평생교육 6진분류 표'의 구성을 기반으로 한다.

데이터의 가용성은 2019년 8월 25일 기준으로 인터넷 사이트 조사기관인 '인터넷트렌드'의 내용에 따르면 Naver 점유율은 55.30%로 1위이고 이어서 Google(37.16%), Daum(6.07%), Zum(0.57%), Bing(0.48%) 순으로 나타났다. 즉, Google과 Naver 데이터를 활용하면 국내 검색엔진 사용자 중 92.46%의 데이터를 수집·활용할 수 있다 [7][11].

4) 문제 해결을 위한 논리적 모델 설계

API를 활용한 수집과 시각화를 위한 적합한 프로그래밍 언어 및 프로그램을 선정하고, 교육 특성을 고려한 논리적 모델을 설계한다.

5) 과제 타당성 평가

수집된 데이터의 연령, 학력, 성별 등을 고려하여 지역별 특성을 반영하였는지 적합성을 평가한다.

6) 과제 확정 및 분석 계획 수립

전체 과정에서 프로토타이핑 빅데이터 분석 기획 과정을 적용하여 가능성 검토 후 과제 확정 및 분석 계획을 수립한다.

2. 데이터 수집 과정

Naver DataLap과 Google trends 모두 검색어 트렌드 API를 개발하여 제공하고 있으며 Python, PHP, Node.js, Java, Rstudio, c# 등을 활용할 수 있다. API구현은 프로그래밍 내용은 GitHub를 통해 해당 프로젝트를 참조 URL주소:<https://github.com/Computer-x64/Kim-Hyeon-Seong.git>

3. 지역 맞춤형 교육프로그램 선정 모형 개발 표본

연구방법과 이론적 배경에 대한 부분을 충분히 이해하는 것을 선결과제로 하지만, 선결과제의 어려움 및 비전공자를 고려하여 실제 프로그래밍한 내용과 개별내용에 대한 각주를 포함하고 순차적으로 세부 단계별로 개발모형 표본을 설명한다.

1) Python 및 모듈 설치

Python은 윈도우 윈도우도스창(cmd.exe)에서 다음 명령어를 통해 `curl https://bootstrap.pypa.io/get-pip.py -o get-pip.py(pip install python get-pip.py(python install) 설치하거나 웹주소로 다운로드 할수 있으며, pip와 python다운로드 후에 pytrends와 pandas 모듈을 다음 명령어를 통해 pip install pytrends, pip install pandas 설치한다.`

2) 기반 자료 생성

기반 자료 생성에는 가장 보편적으로 활용되고 있는 ‘평생교육 6진분류 표’의 내용을 기반한 .txt(텍스트파일) 혹은 .csv(엑셀파일)로 데이터를 생성한다. 본 연구자의 경우 6진분류표를 그림 3 과 같이 .csv파일로 정리한 후 그림 4 와 같이 .txt파일에 복사하고 그림 5 와 같이 .txt 저장단계에서 인코딩 방식을 UTF-16으로 하였다.

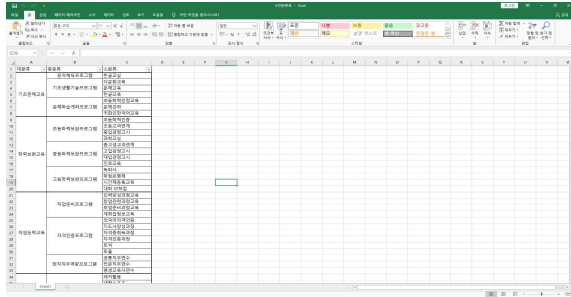


그림 3. 6진 분류표 .csv생성
 Figure 3. Generating a 6-level classification table in .csv format



그림 4. 6진 분류표 .txt 복사
 Figure 4. Copying the 6-level classification table in .txt format

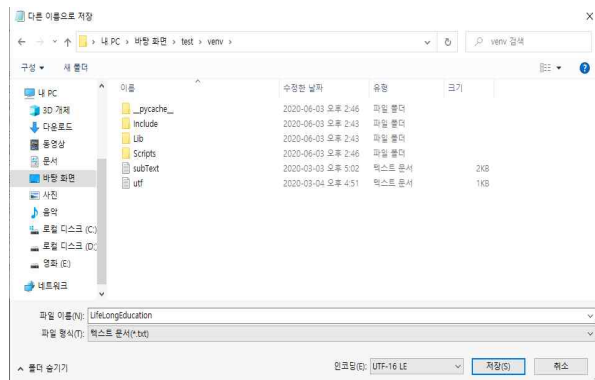


그림 5. 6진 분류표 .txt 인코딩
 Figure 5. Encoding the 6-level classification table in .txt format

3) 데이터 수집

데이터 수집의 표본은 다음 표 4 와 같고 pytrends 데이터 생성과 형태는 다음 그림 6, 7 과 같다.

표 4. 데이터 수집 표본

Table 4. Data collection sample

```
Pytrends API python
from pytrends.request import TrendReq
# Google에 연결pytrends를 사용한다.

import pandas as pd
# pandas 모듈을 가져오고pd라고 입력하여 사용한다.

import csv
# csv모듈을 가져온다.

File = open("LifeLongEducation.txt", "r", encoding='UTF-16')
# LifeLongEducation.txt파일을 읽어오고
인코딩은UTF-16으로하여File에 할당한다.

subGroup = File.readlines()
# subGroup에 인코딩한File을 라인별로 읽어온다.

print(subGroup)
# 읽어온 내용을 출력한다(확인).

pytrends = TrendReq(hl='ko', tz=540,)
# hl='한국', tz=표준시간대,)

for i in range(len(subGroup)//2):
# 두라인씩 읽자 순차적 검색을 실시 현재 구글트렌드 제한상
최대5개까지만 가능

keywords = [subGroup[i * 2]].strip(), subGroup[(i * 2) +
1].strip()]
# 5개씩 할 경우 subGroup[(i*5)].strip(),
subGroup[(i*5)+1].strip(), subGroup[(i*5)+2].strip(),
subGroup[(i*5)+3].strip(), subGroup[(i*5)+4].strip())

pytrends.build_payload(keywords, cat=0,
timeframes='2018-01-01 2018-12-31', geo='KR', gprop='')
# cat은 별도의 범주를 정하는 경우(카테고리), timeframes='자료
시간대' 최근5년인경우today 5-y', gprop='필터링 할 속성'

#getdatainfo = pytrends.interest_over_time() 'today 5-y'
getdatainfo = pytrends.interest_by_region(resolution='CITY',
inc_low_vol=True, inc_geo_code=False)
# resolution='CITY'는 도시수준, 'COUNTRY'는 국가수준, 'DMA'는
메트로수준, 'REGION' 지역수준 데이터반환,
inc_low_vol=(True/False)낮은 볼륨 국가/지역에 대한 데이터 포함,
inc_geo_code=(True/False)데이터의 이름과 함께 국가ISO코드 포함

getdatainfo.to_csv("LLE" + str(i) + ".csv", "w",
encoding="EUC-KR")
# 데이터를 읽어LLE(i).csv(엑셀) 파일에 작성한다. LLE0 ~ LLE43
까지 생성된 데이터를 구분기호를 두하여 CSV로 생성시킬 것
이기 때문에 구분기호를 구분하였을 경우 한줄이
기입되는EUC-KR로 인코딩 (UTF-16로 지정하지만 해당 데이터의 경우
CSV에서 한줄데이터가 순상되는 경우가 발생하였음)
File.close()
# 열었던 파일을 닫는다.
```

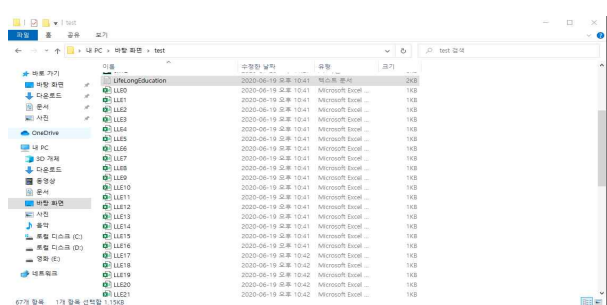


그림 6. pytrends 데이터 생성
 Figure 6. Generating pytrends data

	A	B	C
1	geoName	역사	토익
2	강원도	57	43
3	경기도	58	42
4	경상남도	57	43
5	경상북도	59	41
6	광주광역시	54	46
7	대구광역시	52	48
8	대전광역시	51	49
9	부산광역시	54	46
10	서울특별시	56	44
11	울산광역시	49	51
12	인천광역시	58	42
13	전라남도	61	39
14	전라북도	62	38
15	제주특별자치도	70	30
16	충청남도	60	40
17	충청북도	59	41

그림 7. pytrends 생성된 데이터 형태
 Figure 7. Format of generated pytrends data

4) 데이터 가공

pytrends를 통해 생성된 데이터를 가공하기 위해 첫

번째로 생성된 csv파일을 별도의 폴더를 만들어 집합시키고 아래와 같이 명령프롬프트(CMD)를 통하거나, 확장자를 배치파일(.bat)으로 만들어 세로 기준 병합한다. 가로 기준 병합 시 Microsoft Excel의 VBA(Visual Basic for Applications)의 모듈 삽입을 통해 다음 표 5 와 같이 병합한다.

표 5. 데이터 가공 방법

Table 5. Data processing methods

파일병합	명령프롬프트(CMD)
	<pre>C:\Users\Computer\Desktop\test\all>type *.csv>all.csv LLE0.csv LLE1.csv... # 폴더내 모든 .CSV파일이 합쳐짐(기본 세로병합) # 해당없는 CSV파일 주의</pre>
파일병합	배치파일 생성(.bat)
	<pre>type *.csv>all.csv # 작성후 확장자 .bat으로 저장후 집합시킨 폴더로 이동후 실행 # 해당없는 CSV파일 주의</pre>
세로파일병합	VBA(Visual Basic for Applications)
	<pre>Sub SheetUnit() Dim i As Integer Dim ShtA As Worksheet Dim rngB As Range Application.ScreenUpdating = False Set ShtA = Sheets(1) For i = 2 To Sheets.Count Set rngB = ShtA.Cells(3, Columns.Count).End(1).Offset(-2, 1) Sheets(i).UsedRange.Copy rngB Next i Cells.EntireColumn.AutoFit End Sub # Microsoft Excel 실행 후 Alt+F11키를 눌러 삽입-모듈에 해당 코딩을 한후 F5키를 눌러주면 병합이 된다. # 해당코딩은 가로기준으로 병합시 사용</pre>

이후 교차방식으로 비율의 근하위 상대배수를 구하고 최저값을 절대기준으로 하여 다음 표 6 과 같이 데이터 가공을 마무리 해준다.

표 6. 유효 가공한 데이터

Table 6. Valid processed data

구분	역사 data	토익 data	미술 data	전통 data	독서 data	공예 data	토폴 data	무용 data
교차 비율값	53	65 47	51 35	61 49	70 39	56 30	52 44	48
근하위 상대배수	1.13	1.88	1.04	1.58	2.33	1.27	1.08	1
최저값 절대기준	10.97	9.73	5.24	5.03	3.22	1.38	1.08	1

해당 데이터의 경우 비율의 제한과 절대값으로 데이터 순회형 함수와 연산자에 sorted 연산자와 더불어 교차분석을 추가하거나, Microsoft Excel의 VBA를 통해 교육 6진 분류표 기반 키워드 txt파일을 수정하고 파일 병합

전 파일명 뒤 숫자를 이진수로 다음 표 7 과 같이 변환해야 한다. 이후 그림 8 과 같이 파일 병합을 실시 한다.

표 7. 병합 전 파일명 이진수 계산식

Table 7. Binary calculation formula for file names before merging

구분	내용
이진수변환	<pre> 2 42 0 2 21 1 2 10 0 2 5 1 2 2 0 1 1 1 ∴ 101010 (생성된 파일 수가 42개인 경우) </pre>

데이터 시각화는 다양한 프로그램들이 존재하지만 가장 쉽게 활용 가능하며 무료로 제공되고 있는 빅데이터 센터(kbig.kr)에서 그림 8 과 같이 실시하였다.



그림 8. 빅데이터 시각화 디자인 조정

Figure 8. Adjusting the design of big data visualization

IV. 토론

교육기관의 교육사들은 프로그램 개발 방법으로 주로 전문 강사들에게 프로그램 커리큘럼을 제공받는 것을 채택한다. 지역의 교육 전문가와 기관 내 담당자가 커리큘럼을 만든 강사는 면접을 통해 평가하고 커리큘럼이 결정되면 강사는 홍보를 통해 학습자를 모집한다. 만약 정원을 채우지 못한다면 폐강하게 되는데, 이 경우 계획된 시간과 시설의 활용도는 낮아지는 문제를 정의하였다. 이런 문제를 해결하기 위하여서는 교육이 지역사회와 함께 성장하고, 주민 주도적인 방식으로 교육프로그램을 선정해야 한다. 즉, 교육 문화를 정착시키기 위해 교육프로그램 개발 및 선정의 과정은 지역사회 주민의 욕구를 충분히 반영하고 함께 성장 가능한 개발 모형으로 빅데이터를 활용한 교육프로그램 개발 모형을 제시했다.

현재 빅데이터 연구는 공공, 금융, 기업 경영, 통신·미디어 분야 등 다양한 분야에서 인정받고 있는 것에 비해 교육 분야에서의 연구나 활용은 미비한 실정이다. 이에 본 연구자는 문헌 고찰을 통해 교육 및 빅데이터의 개념과 특성 그리고 기술 및 연구 활용 등의 이론을 교육 분야에 활용을 위한 선결과제와 연구의 기초자료로 제공하

였고, 교육 데이터 수집 방법과 교육의 특성에 적절한 빅데이터 활용 방법을 제시하고 이를 기반으로 하여 한 지역 맞춤형 프로그램 선정 모형을 개발하였다. 개발 과정에서 발생한 문제는 세 가지로 첫째, 데이터 가용성 편성단계 가운데 Google Trends는 나이, 연령, 성별이 고려되지 않았으며 Naver DataLap은 지역별 특성이 고려되지 못하였기 때문에 지역 특성에 대한 부분을 우선시하여 Google Trends를 편성하였다. 둘째, 문제 해결을 위한 논리적 모델 설계단계 중 코딩 언어 선택의 문제가 있었고, 빅데이터를 분석하기 위하여 다양한 모듈러 기능을 제공하면서 비전공자 또한 비교적 단순하고 배우기 쉬우며 확장성이 좋은 python 선정하였다. 셋째, 데이터 수집·가공 중 절대값 및 검색어 개수 제한을 기준으로 한 비율표기로 인한 분석값의 오류가 있었고 최저값 및 반복 합수를 기준으로 절대값 및 교차방식 정렬을 프로그래밍(windows cmd, VBA, python)하여 해결하였다.

본 연구에서 제시한 맞춤형 교육프로그램 모델은 실질적 활용 면에 있어, 국가승인통계인 ‘평생학습 개인 실태조사’ 처럼 1년 후에 분석하지 않고 실시간으로 데이터가 제공되는 방식으로 활용 부분에 있어서도 선택적 분석이나 미래예측 등 자유도가 매우 높아 교육 분야에 빅데이터가 충분한 필요성과 가치가 있음을 알 수 있다. 뿐만 아니라 표본 모형에 사용되고 있는 모든 프로그램은 무료로 제공되고 있으며, 프로그래밍 특성상 커뮤니티 또한 활발하게 교류가 이루어지고 있어 추후 수정 및 보완 시에도 매우 용이하여 더욱 완성도 높은 교육프로그램 개발 모형을 개발할 수 있다.

V. 결론

프로그램 선정 모형 개발 과정은 총 6단계로 구분하여 설명할 수 있다.

첫 번째, 문제 발굴 및 정의단계에서 현재의 교육프로그램 운영방식의 시설 활용도가 낮아지는 문제를 제기하였다. 두 번째, 문제 해결을 위한 개념적 대안 설계단계에서 교육 전문 강사진들에게 프로그램 교육과정을 제공받지 않고 주민의 욕구를 충분히 반영한 교육프로그램 개발 방법을 제시하였다. 세 번째, 데이터 가용성 편성단계에서 Google Trends와 Naver DataLap 활용하여 크롤링(-트래킹)기법으로 데이터를 수집하며, 내용은 ‘교육

6진 분류표’의 내용을 토대로 하였다. 네 번째, 문제 해결을 위한 논리적 모델 설계단계에서 API를 활용한 데이터 수집과 시각화를 위한 적절한 프로그래밍 언어 및 프로그램을 선정하고, 교육 특성에 적절한 논리적 모델을 설계하였다. 다섯 번째로는, 수집된 데이터의 연령, 학력, 성별 등을 고려한 지역별 특성이 반영되었는지 적합성을 평가하고 마지막으로 전체 단계에서 프로토타이핑 빅데이터 분석 기획 과정을 적용하여 가능성을 검토 후 과제 확정 및 분석 계획을 수립하도록 하였다.

References

- [1] Big Data Center [Website]. “Retrieved from kbig.kr”, October 2, 2019.
- [2] Caffarella, R. S. “Planning programs for adult learners: A practical guide for educators, trainers, and staff developers (2nd ed.)”, San Francisco, CA: Jossey-Bass. Retrieved from, 2002.
- [3] CDC; Flu Near you [Website]. “Retrieved from www.google.org/flutrends”, November 11, 2019.
- [4] Cervero, R. M. & Wilson, A. L. “The song remains the same: The selective tradition of technical rationality in adult education program planning theory. International Journal of Lifelong Education”, 16(2), pp. 84-108, 1997.
- [5] Cho, D. Y. “Trends in adult education program development theory in the United States: 1990-2001. Journal of Lifelong Education”, 10(1), pp. 27-42, 2004.
- [6] Data Store [Website]. “Retrieved from data.store.or.kr”, October 28, 2019.
- [7] Google Trends [Website]. “Retrieved from trends.google.co.kr”, November 18, 2019.
- [8] IDC(International Data Corporation) [Website]. “Retrieved from https://www.idc.com/”, November 14, 2019.
- [9] Kim, J. H. “Theory of developing lifelong education programs.”, Seoul: Kyoyook-Kwahak-Sa, 2003.
- [10]Microsoft Excel [Website]. “Retrieved from https://products.office.com/ko-kr/excel” October 25, 2019.
- [11]Naver DataLab [Website]. “Retrieved from datalab.naver.com”, November 6, 2019.
- [12]R Project. “Retrieved from https://www.r-project.org/”, October 21, 2019.
- [13]RStudio. “Retrieved from https://rstudio.com/”,

October 30, 2019.

- [16] Statistical Geographic Information Service. "Retrieved from sgis.kostat.go.kr", October 12, 2019.
- [17] Warren, R. M. "Program planning and development in adult education: Where we are at the beginning of the 21st century. 2003 Hawaii International Conference on Education", 2003..
- [18] Yang, B. & Cervero, R. M. "Power and influence styles in programme planning: Relationship with organizational political context. International Journal of Lifelong Education", 20(4), pp. 289-296, 2001.