

2015 개정 중학교 정보 교과서의 ‘문제 해결과 프로그래밍’ 단원 핵심 개념 분석

최현종

서원대학교 컴퓨터교육과 교수

Analysis of Core Concepts in Problem Solving and Programming Unit of Informatics Subject Textbooks in Middle School Revised in 2015

Hyun-Jong Choe

Professor, Department of Computer Education, Seowon University, Chungbuk 28674, Korea

[요 약]

본 연구는 2015 개정 교육과정의 중학교 정보 교육과정과 세 가지 종류의 정보 교과서의 문제 해결과 프로그래밍 단원 관련 내용을 대상으로 텍스트 마이닝 방법으로 핵심 개념을 비교 분석하였다. 중학교 정보 교육과정과 교과서의 단원을 워드 클라우드와 단어 빈도 분석을 통해 핵심 개념을 추출하였고, 교육과정과 교과서간의 핵심 개념에 대한 분석을 통해 교과서에서 강조하고 있는 개념들을 확인하였다. 또한 교육과정과 교과서의 문장에 포함된 단어들의 연관 분석을 통해 핵심 개념이 어떻게 서로 연관되어 사용되고 있는지도 확인하였다. 분석 결과 교육과정에서 제시된 핵심 개념과 연관 개념들이 세 가지 종류의 교과서에서도 비슷한 중요도로 사용되고 있었지만, 교과서에 따라 핵심 개념과 연관 개념들이 서로 조금씩 상이하다는 것을 발견하였다.

[Abstract]

This study compared and analyzed core concepts in the problem solving and programming unit of the 2015 revised middle school Informatics subject curriculum and three textbooks using text mining techniques. The core concepts of the unit in the 2015 revised middle school Informatics subject curriculum and three textbooks were extracted through word cloud and word frequency analysis. Through the analysis of the core concepts between the curriculum and the textbook, the concepts emphasized in the textbook were identified. In addition, the association analysis of words included in the curriculum and textbook sentences confirmed how the core concepts are related to each other. The analysis result showed that the core concepts and related concepts presented in the curriculum were used with similar importance in three textbooks. And it was also founded that the core concepts and related concepts were differed slightly from each other.

색인어 : 2015 개정 교육과정, 연관 분석, 빈도 분석, 정보 교과, 교과서 분석

Key word : 2015 revised curriculum, Association analysis, Frequency analysis, Informatics subject, Textbook analysis

<http://dx.doi.org/10.9728/dcs.2020.21.1.63>



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.

Received 06 November 2019; Revised 31 December 2019

Accepted 23 January 2020

*Corresponding Author; Hyun-Jong Choe

Tel: +82-43-299-8432

E-mail: blueland@seowon.ac.kr

1. 서론

2015 개정 교육과정의 고시로 초등학교와 중학교에서 소프트웨어 교육이 필수화되었고, 2018년부터 전국의 중학교에서 정보(Informatics) 과목이 필수로 가르쳐지고 있다[1][2]. 기존의 ICT 활용 교육에서 탈피하여 컴퓨터 과학 교육과 컴퓨팅 사고력(Computational Thinking) 교육을 강조하고 있는 정보 과목이 실제 중학교에서 어떤 내용이 어떻게 가르쳐지고 있는지 점검하는 것은 새롭게 필수가 된 정보 과목에는 중요한 문제이다 [3]. 구체적으로 어떤 내용을 어떻게 가르치고 있는지를 확인하는 가장 좋은 방법은 바로 정보 교과서를 분석하는 방법이다. 교과서는 국가에서 고시한 교육과정을 구체적으로 실현한 수업 자료로, 교사는 교육과정에서 구현하고자 하는 목표와 내용을 교과서로 학생을 지도하기 때문이다.

교과서 분석 연구는 교육 현장을 대변하는 가장 중요한 기본적인 교과 교육 연구이다. 이에 많은 연구자들이 새로운 교육과정이 고시되어 교과서가 출판되면 교과서 분석 연구를 실시하였다. 기존의 교과서 분석 연구는 크게 정량적 분석 연구와 정성적 분석 연구로 구분할 수 있다[4]. 정량적 연구의 대표적 경우는 교과서의 내용에 대한 롬니 지수(Romey Index)를 측정하는 연구인데, 롬니 지수는 교과서에 제시된 내용이 얼마나 탐구적이나를 측정하는 지수로 연구자가 교과서의 본문 내용을 사실, 결론, 질문 등으로 구분하여 이에 관한 지수를 산출한다[5]. 정성적 연구의 경우는 내용 분석법으로 교과서의 내용이나 형식 등을 연구자의 특정 기준으로 분석하는 방법이다.

교과서 연구 방법인 롬니 지수 산출법과 내용 분석법은 연구자의 가치 판별과 해석이 포함된 방법이다. 이에 본 연구는 객관적 방법으로 교과서의 내용을 분석할 수 있는 텍스트 마이닝 기법을 사용하고자 한다. 빅데이터 연구가 일상생활과 다양한 학문 분야에서 진행되고 있는데, 교육 분야에서도 선도적으로 적용되는 시기가 도래하였다. 특히, 교과서는 텍스트, 그림과 사진 등 다양한 정보가 담겨져 있는 수업의 기본 자료로 교과서에 포함된 텍스트 정보 자체를 취합하여 분석하는 연구는 이전에는 매우 어려웠지만, 현재는 교과서가 디지털 콘텐츠로 보급되면서 다양한 텍스트 마이닝 기법을 이용한 연구가 이루어질 수 있다. 따라서, 본 연구는 2015 개정 교육과정의 중학교 정보 교육과정에서 가장 중요하게 다루어지고 있으며, 가장 많은 시수를 차지하고 있는 문제 해결과 프로그래밍 단원의 내용과 세 종류의 중학교 정보 교과서 내용에 포함된 텍스트를 분석하는 것을 목적으로 한다.

연구에 적용된 분석 방법은 단어 빈도 분석과 연관 분석을 실시하였다. 빈도 분석은 단순한 방법이지만 교과서의 내용 중에서 가장 많이 빈번하게 제시되고 있는 단어를 추출하는 기법으로, 이 방법으로 교과서에서 중요하게 다루어지는 단어, 즉 핵심 개념을 파악할 수 있을 것이다. 또한 추출된 단어 간의 연관 관계 분석을 통해 핵심 개념간의 연관성도 추측할 수 있기 때문에 교과서 연구 중에서 가장 기본적이면서도 객관적인 자

료로 활용될 수 있을 것이다.

II. 관련 연구

2-1 교과서의 정량적 분석 연구

교과서 분석 연구 방법은 정량적 분석과 정성적 분석으로 구분될 수 있는데, 지금까지 교과서 연구의 대부분은 정량적 분석 방법을 사용하였다. 이 정량적 분석 방법에는 롬니 지수를 측정하는 방법이 있는데, 롬니 지수는 교과서의 내용이 학습자의 탐구 활동을 얼마나 유도하는지를 정량적으로 측정할 수치이다. 강오한(2019)은 2015 개정 교육과정에 의해 출판된 중학교와 고등학교 정보 교과서 네 종류를 선정하여 롬니 지수를 산출하였다[6][7]. 또한 2009 개정 교육과정에 의해 출판된 중학교와 고등학교 정보 과목의 교과서도 롬니 지수로 교과서의 탐구성을 연구한 사례도 있다[8]-[10].

교과서의 정량적 분석 방법 중의 하나인 롬니 지수 분석은 다음과 같은 두 가지의 단점을 가지고 있다[11]. 첫 번째로 롬니 지수는 탐구성이란 한 가지 경향을 분석하는 방법이라는 것과 두 번째는 롬니 지수를 계산하기 위해 연구자는 교과서에 수록된 본문, 자료, 활동, 평가 등이 사실인지, 혹은 정의, 결론인지를 판단해야 한다. 또한 교과서의 질문이 즉시 답을 요구하는 것인지, 자료 분석을 요구하는 것인지, 자신의 결론을 만들게 하는지, 혹은 분석하도록 하는지를 구분해야 한다는 것이다. 이 과정은 교과서의 내용 중에서 명백하게 구분되는 내용도 있지만, 그렇지 못한 내용도 있기 때문에 연구자는 임의로 구분지어 분류해야 하는 과정이 필요하다. 따라서 롬니 지수 분석법은 경우에 따라 연구자의 가치 판별과 해석이 포함될 수 있다.

2-2 교과서의 정성적 분석 연구

교과서의 정성적 분석 방법은 교과서의 내용이나 체제, 형식을 연구자가 정한 판단 기준으로 분석하는 연구이다. 김정량(2019)은 필수화된 초등학교의 소프트웨어 교육이 실과 과목의 교과서에 어떻게 적용되었는지를 확인하기 위해 교육과정에 제시된 성취기준을 기준으로 교과서에 수록된 내용을 분석하였다[12]. 김수환(2018)은 중학교 정보 과목 교과서에 포함된 정의, 개념, 동기 유발, 활동 및 예시 등이 교육과정의 추상화에 얼마나 부합하는지를 연구하였다[13]. 최현중(2014)은 앤더슨의 신교육목표 분류법으로 중학교 정보 교과서에 수록된 학습 목표를 분석하였다[14]. 이렇듯 정성적 분석 방법은 연구자가 정한 연구의 주제에 따라 분석 기준을 적용한 연구들이 진행되고 있다.

2-3 교육의 빅데이터 관련 연구

정보통신 기술이 발달되면서 대부분의 연구 자료가 디지털

화되고, 이런 디지털 콘텐츠들이 인터넷을 통해 공유되면서 빅데이터(Big Data)라는 새로운 학문 분야가 등장하였고, 다양한 학문 분야에서 빅데이터를 활용한 연구들이 진행되고 있다. 교육 분야에서도 다양한 연구들이 진행되고 있는데, 권선아, 김한나, 이수영 외(2019)는 위크숍과 신문기사로부터 미래 교육 관련 텍스트를 텍스트 마이닝으로 분석하여 핵심 내용을 제시하는 연구를 하였다[15]. 신중호와 최재원(2019)은 한 대학에서 9 학기 동안 실시되었던 서술형 형태의 강의 평가 내용을 텍스트 마이닝으로 분석하는 연구를 진행하였고[16], 고수정(2018)은 중학생들을 학습 습관으로 그룹으로 묶어, 그 특성으로 빅데이터 군집 분석으로 학습 성취도를 연구하였다[17]. 특히, 김셋별과 채정현(2019)은 2015 개정 교육과정에 의해 출판된 가정 교과서의 생활 문화라는 단원의 핵심 개념을 텍스트 마이닝으로 분석하는 연구를 수행했고[18], 최현중(2019)은 고등학교 교과서 일부를 대상으로 텍스트 정보를 대상으로 텍스트 마이닝 분석을 실시하였다[11]. 빅데이터 관련 연구에서 텍스트 마이닝을 사용한 연구들이 특히 많은데, 이는 텍스트 마이닝이 비구조화된 디지털 텍스트에서 의미있는 정보를 추출하여 새로운 지식을 발견하는 방법으로 교과서나 신문기사, 서술형 문장 등에서 유용하게 사용할 수 있는 방법이기 때문이다. 이에 본 연구는 교과서안에 수록된 디지털 텍스트를 텍스트 마이닝으로 분석하여, 중학교 정보 교육과정과 교과서의 핵심 개념을 비교하고자 한다.

III. 연구의 설계

3-1 연구의 대상

본 연구는 2015 개정 교육과정의 중학교 정보 교육과정에서 문제 해결과 프로그래밍 단원에 제시된 내용과 이에 맞춰 출판된 세 가지 종류의 중학교 정보 교과서에서 관련 내용을 분석하는 것을 목적으로 한다. 중학교 정보 교과서는 모두 17개 출판사의 교과서가 있는데, 본 연구에서는 샘플 데이터로 연구자 소재 지역의 중학교에서 사용하고 있는 교과서 세 가지 종류를 선택하여 분석 대상으로 하였고, 분석 도구는 R 프로그램을 사용하였다.

3-2 연구의 절차 및 방법

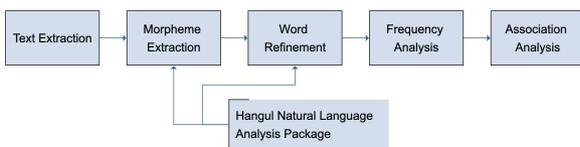


그림 1. 연구의 절차
Fig. 1. Process of study

연구의 절차는 그림 1과 같은데, R 프로그램으로 분석 결과를 도출하는 과정이다. 먼저 교육과정과 중학교 정보 교과서의 디지털 콘텐츠 파일(PDF)에서 텍스트를 먼저 추출하고, 이를 한글 자연어 분석 패키지를 사용하여 형태소 분석 가능 형태로 변환하여 추출하였다. 형태소 형태로 구분된 단어들은 정교한 분석을 위해 정선하는 과정을 거치게 되는데, 한글 자연어 분석 패키지와 한글 사전 패키지를 이용하여 무의미한 문자들을 제외시키면서 단어를 정제하였다. 네 번째 과정은 추출된 단어들을 빈도 분석하는 과정으로 시각화 패키지를 이용하여 워드 클라우드로 표현해 보고, 가장 많이 사용된 단어들을 추출해 보는 빈도 분석을 실시하였다. 마지막으로 문장 단위로 사용되는 단어들의 연관성을 분석하기 위해 연관 분석을 실시하였다. 연관 분석은 동시에 발생하는 사건들의 규칙을 수치화하여 연관 규칙을 표현하는데, 한 사건과 다른 사건 사이의 연관성을 찾아낼 때 사용한다[19]. 데이터 사이의 연관성을 찾기 위해 사용되는 대표적인 지표에는 지지도(support)와 신뢰도(confidence)가 있다.

표 1. 연관 분석의 측정 지표

Table 1. Performance measures in association analysis

Measures	Numerical expression
Support	$P(A \cap B)$
Confidence	$P(A B) = \frac{P(A \cap B)}{P(B)}$

표1에서 제시하고 있는 지표를 이용하여 두 데이터 A, B의 연관 정도를 수치로 보여 줄 수 있다. 예를 들어, 지지도는 문장에서 두 개의 단어 A, B를 모두 가지고 있는 확률을 의미하고, 신뢰도는 조건부 확률로 단어 A가 포함된 문장 중에서 A와 B, 둘 다 가지고 있는 확률을 의미한다.

IV. 연구의 결과

4-1 빈도 분석 결과

2015 개정 중학교 정보 교육과정의 문제 해결과 프로그래밍 관련 내용과 세 종류의 중학교 정보 교과서의 문제 해결과 프로그래밍 단원의 콘텐츠 내용을 텍스트 마이닝으로 분석하여, 이를 워드 클라우드로 표현하면 그림 2와 같다.

워드 클라우드로 표현된 단어를 살펴보면 선별된 단어와 그 중요도가 서로 상이하다는 것을 확인할 수 있다. 즉 동일한 단어가 추출되어 표시되더라도, 교육과정과 교과서에 제시된 빈도와 비율이 각기 다르다는 것을 확인할 수 있다.

표 3. B, C 교과서의 단어 빈도 분석

Table 3. Word frequency analysis of B and C textbooks

Rank	B (Total N=1030)			C (Total N=790)		
	Word	Fre. (N)	Per. (%)	Word	Fre. (N)	Per. (%)
1	Problem	56	5.4	Problem	37	4.7
2	Solve	40	3.9	Solve	19	2.4
3	Algorithms	38	3.7	Do	18	2.3
4	Condition	24	2.3	Language	17	2.2
5	Program	24	2.3	Structure	16	2.0
6	Use	19	1.8	Use	15	1.9
7	Do	17	1.7	Sequential	15	1.9
8	Structure	16	1.6	Program	15	1.9
9	Language	16	1.6	Process	14	1.8
10	Variable	15	1.5	Algorithms	14	1.8
11	Applicate	15	1.5	Variable	12	1.5
12	Sequential	14	1.4	Data	12	1.5
13	Result	14	1.4	Represent	12	1.5
14	Loop	14	1.4	Loop	11	1.4
15	Process	13	1.3	Method	11	1.4
16	Computer	13	1.3	Input	11	1.4
17	Programming	13	1.3	Programming	11	1.4
18	Method	12	1.2	Result	10	1.3
19	Process	12	1.2	Output	10	1.3
20	Select	12	1.2	Computer	10	1.3

4-2 연관 분석 결과

중학교 정보 교육과정과 세 교과서의 문제 해결과 프로그래밍 단원에 기술된 문장 안의 단어 연관 관계를 살펴보기 위해 연관 분석을 실시하였다.

표 4는 중학교 정보 교육과정과 세 교과서의 문제 해결과 프로그래밍 단원의 문장에서 동시에 출현하는 가능한 모든 쌍의 단어 간 연관 규칙을 발췌하여 지면상 그 일부를 정리한 표이다. 그림 3에 제시된 연관 그래프는 문장 안에 포함된 단어들의 연관도를 먼저 분석한 후, 이를 그래프로 표현한 것이다.

표 4. 교육과정과 세 교과서의 최상위 연관 규칙

Table 4. The top association rules of curriculum and three textbooks

	Left Word	Right Word	Supp.	Conf.	Fri. (N)
Curriculum	{Problem}	{Solve}	0.56	0.85	29
	{Problem, Solve}	{Process}	0.27	0.48	14
A	{Problem}	{Solve}	0.18	0.84	33
	{Problem, Solve}	{Do}	0.06	0.36	12
B	{Solve}	{Process}	0.09	0.37	11
	{Problem, Solve}	{Process}	0.09	0.37	11
C	{Problem}	{Abstraction}	0.06	0.22	7
	{Problem, Solve}	{Method}	0.05	0.33	6

표 4에서 정보 교육과정은 문제, 해결이란 단어가 동시에 발견될 확률인 지지도는 0.56 정도이지만, 문제라는 단어가 포함된 문장에서 해결이란 단어가 발견될 확률인 신뢰도는 0.85로 매우 높게 나타난 것을 확인할 수 있다. A 교과서도 {문제}-{해결}의 지지도는 0.18에 불과했지만, 신뢰도는 0.84로 매우 높게 나타난 것을 확인하였는데, 이와 같은 방식으로 다른 주제어들 간의 쌍들도 연관 관계를 해석할 수 있다.

중학교 정보 교육과정에 포함된 문제해결과 프로그래밍 단원 문장안의 단어들에 관한 연관도는 지지도 0.15, 신뢰도 0.3 수준으로 계산해 보니 모두 61개의 규칙을 발견하였다. 세 종류의 교과서는 지지도 0.05, 신뢰도 0.01 수준으로 A 교과서는 40개의 규칙, B 교과서는 120개의 규칙, C 교과서는 31개의 규칙이 발견되었다. 연관 규칙의 빈도수를 살펴보면 B 교과서가 다른 두 종류의 교과서보다 훨씬 많은 연관 규칙을 가지고 있는데, 이는 서로 연관성 있는 단어들을 포함한 문장의 빈도수가 확률적으로 높다는 것이다. 즉, 다른 교과서보다 서로 연관된 단어들이 많이 진술되어 있어 교과서에서 두 개의 단어를 연관 지어 강조하고 있다는 의미이다.

또한 표 4의 연관 규칙 중에서 가장 높은 빈도를 보인 두 개의 연관 단어들은 A 교과서는 {문제}-{해결}, B 교과서는 {해결}-{과정}, C 교과서는 {문제}-{추상화}이다. A와 B 교과서는 문제의 해결과 해결 과정에 초점을 맞추어 문장을 진술한 횟수가 가장 많은 것에 비해 C 교과서는 문제의 추상화 관점에서 진술된 문장의 개수가 가장 많은 것을 확인할 수 있었다. 세 개의 연관 단어들은 A 교과서가 {문제, 해결}-{하다}, B 교과서는 {문제, 해결}-{과정}, C 교과서는 {문제, 해결}-{방법}이다. 세 개의 연관 단어들은 세 종류의 교과서 모두 문제의 해결 관점을 강조하여 문장을 서술하고 있는 것을 확인할 수 있었다.

집필할 때 좋은 참고 자료가 될 것으로 기대된다.

참고문헌

- [1] Ministry of Education, and Ministry of Science, ICT and Future Planing, Plan for human resource development for SW-oriented society, 2015.
- [2] T. W. Lee and H. J. Choe, Informatics Education, Hanbit Express, 2016.
- [3] H. N. Jung and Y. S. Jhun, "Analysis of Instruction Materials for Physical Computing from the Perspective of Computer and Information Literacy", *Journal of The Korean Association of information Education*, Vol.22, No.4, pp.473-489, Dec. 2018.
- [4] O. H. Kang and J. M. Park, "Analysis of 'Basic Information Technology' Textbooks in Technical High Schools", *The Journal of the Korea Contents Association*, Vol.13, No.3, pp.485-494, March 2013.
- [5] B. G. You, J. M. Kim and W. G. Lee, "Implication for Construction Computing System Unit of the 2015 Revised Curriculum", *The Journal of Korean association of computer education*, Vol.19, No.2, pp.31-40, June 2016.
- [6] O. H. Kang, "The Analysis of Inquisitive Tendency in Middle School Informatics Textbooks Based on the 2015 Revised National Curriculum", *The Journal of Korean association of computer education*, Vol.22, No.1, pp.21-28, March 2019.
- [7] O. H. Kang, "Analysis on the Aspect of Learners' Explorations Reflected in High School Informatics Textbooks Based on the 2015 Revised National Curriculum", *KIPS Transactions on Computer and Communication Systems*, Vol.8, No.5, pp.103-110, May 2019.
- [8] O. H. Kang and B. S. Kim, "The Analysis of Inquisitive Tendency of Problem-Solving Method and Process Section in High School Informatics Textbooks", *The Journal of Korean association of computer education*, Vol.15, No.4, pp.51-59, Dec. 2012.
- [9] J. M. Kim, et al., "Analysis of Inquiry Tendency in the 'Information Representation and Management' Sections in Middle School Informatics Textbooks", *The KIPS Transactions : Part A*, Vol.19, No.1, pp.9-16, March 2012.
- [10] J. H. Jung and Y. S. Kim, "A study on Analysis of Inquisitive Tendency of 2009 Revised Middle School Informatics Textbooks", *Journal of Next-generation Convergence Information Services Technology*, Vol.3, No.1, pp. 1-8, March 2014.
- [11] H. J. Choe, "Frequency Analysis of Core Concepts of High School Informatics Subject Textbooks Using Text Mining", *Asia-pacific Journal of Multimedia Services Convergent with Art, Humanities, and Sociology*, Vol.9, No.11, pp.419-429, Nov. 2019.
- [12] J. R. Kim, "Analyzing contents of software education area shown in the 2015 revised elementary Practical Art textbook", *Journal of The Korean Association of Information Education*, Vol.23, No.1, pp.9-18, March 2019.
- [13] S. H. Kim, "Analysis of Abstraction Contents in Informatics Textbooks of Middle School According to 2015 Revised Curriculum", *The Journal of Korean association of computer education*, Vol.21, No.5, pp.1-10, Sept. 2018.
- [14] H. J. Choe, "Study of Analysis about Learning Objectives of Informatics Textbooks in Middle School using Anderson's Taxonomy of Educational Objectives", *The Journal of Korean association of computer education*, Vol.17, No.1, pp.51-63, March 2014.
- [15] K. Sun, H. Kim and S. Y. Lee, "Analyzing the Keywords of Future Education using Text-mining Methodology", *Asia-pacific Journal of Multimedia Services Convergent with Art, Humanities, and Sociology*, Vol.9, No.5, pp.11-23, May 2019.
- [16] J. H. Shin and J. W. Choi, "Text mining Analysis of College Students' Descriptive Course Evaluation", *Journal of Learner-Centered Curriculum and Instruction*, Vol.19, No.16, pp.77-99, Oct. 2019.
- [17] S. J. Ko, "Predicting Learning Achievement Using Big Data Cluster Analysis - Focusing on Longitudinal Study", *Journal of Digital Contents Society*, Vol.19, No.9, pp.1769-1778, Sept. 2018.
- [18] S. B. Kim, S. B. Chae and H. Jung, "Analysis of core area of living culture in the high school Home Economics Textbooks Revised in 2015", *Journal of Learner-Centered Curriculum and Instruction*, Vol.19, No.7, pp.871-903, June 2019.
- [19] S. G. Cho and S. B. Kim, "Finding Meaningful Pattern of Key Words in IIE Transactions Using Text Mining", *Journal of the Korean Institute of Industrial Engineers*, Vol.38, No.1, pp.67-73, March 2012.



최현종(Hyun-Jong Choe)

2001년 : 한국교원대학교 대학원 (교육학석사)

2005년 : 한국교원대학교 대학원 (교육학박사-컴퓨터교육)

2005년~2006년: 서원대학교 교수학습센터 연구원

2006년~현 재: 서원대학교 컴퓨터교육과 교수

※ 관심분야 : 컴퓨터교육, 교육과정, 빅데이터 등