



과학교육 분야 자연어 처리 기법의 연구동향 분석

전철홍, 유선아*

한국교원대학교

An Analysis of Trends in Natural Language Processing Research in the Field of Science Education

Cheolhong Jeon, Suna Ryu*

Korea National University of Education

ARTICLE INFO

Article history:

Received 24 October 2023

Received in revised form

11 November 2023

Accepted 9 January 2024

Keywords:

natural language processing,
science education,
role of the teacher

ABSTRACT

This study aimed to examine research trends related to Natural Language Processing (NLP) in science education by analyzing 37 domestic and international documents that utilized NLP techniques in the field of science education from 2011 to September 2023. In particular, the study systematically analyzed the content, focusing on the main application areas of NLP techniques in science education, the role of teachers when utilizing NLP techniques, and a comparison of domestic and international perspectives. The analysis results are as follows: Firstly, it was confirmed that NLP techniques are significantly utilized in formative assessment, automatic scoring, literature review and classification, and pattern extraction in science education. Utilizing NLP in formative assessment allows for real-time analysis of students' learning processes and comprehension, reducing the burden on teachers' lessons and providing accurate, effective feedback to students. In automatic scoring, it contributes to the rapid and precise evaluation of students' responses. In literature review and classification using NLP, it helps to effectively analyze the topics and trends of research related to science education and student reports. It also helps to set future research directions. Utilizing NLP techniques in pattern extraction allows for effective analysis of commonalities or patterns in students' thoughts and responses. Secondly, the introduction of NLP techniques in science education has expanded the role of teachers from mere transmitters of knowledge to leaders who support and facilitate students' learning, requiring teachers to continuously develop their expertise. Thirdly, as domestic research on NLP is focused on literature review and classification, it is necessary to create an environment conducive to the easy collection of text data to diversify NLP research in Korea. Based on these analysis results, the study discussed ways to utilize NLP techniques in science education.

1. 서론

21세기 교육은 디지털화와 기술의 발전에 따라 변화하고 있으며, 이 중심에는 인공 지능(Artificial Intelligence, AI)의 적용이 있다. AI는 다양한 데이터를 처리하여 교육에 혁신을 가져오며(Shaik *et al.*, 2022), 효과적인 교육 전략을 수립할 수 있게 한다. 특히 OpenAI의 ChatGPT와 같은 생성형 AI는 현실적인 대화 능력으로 교육 분야의 패러다임 변화를 주도하고 있다(Tlili *et al.*, 2023). 생성형 인공지능이 가지는 다양한 가능성과 활용은 교사의 업무에 대한 부담을 줄이고 학생들에게 개인화된 학습 경험을 제공할 수 있다(Jho, 2023).

자연어(Natural Language)는 사람들이 일상생활에서 자연스럽게 사용하는 언어를 의미하는데, 컴퓨터가 자연어를 인식하게 하기 위해서는 정해진 프로그래밍 언어로 자연어를 정제하거나 형식에 맞춰 재작성할 필요가 있다(Lim, 2019). 이렇게 컴퓨터가 자연어를 이해하거나 생성할 수 있도록 하는 것을 자연어 처리(Natural Language Processing, NLP)라고 한다. NLP는 AI의 한 분야로서, 인간의 언어를 이해하고 처리하는 능력을 갖추기 위해 개발된 분야이다.

이 NLP 기법은 최근 교육 분야에서 특히 중요한 역할을 하고 있으며, 그 역사 또한 50년 이상 되었다. 초기에는 학생들의 텍스트를 자동으로 채점하거나 대화형 튜터링 시스템을 개발하는 데 중점을 두었으나, 최근의 기술 발전과 함께 교육 분야에서의 활용 가능성이 크게 확장되고 있다(Litman, 2016). 우리나라를 비롯한 외국의 연구 기관과 각급 학교에서는 NLP 기법을 개인화된 학습 관리 시스템, 교사 교육 강화, 언어 평가, 그리고 학생들의 글쓰기 오류 감지와 수정 등 다양한 응용 분야에서 활용하고 있다(Michaud, & McCoy, 2006; Shaik *et al.*, 2022). NLP 기법은 Python, Java, R과 같은 다양한 프로그래밍 언어를 이용한다(Jockers, & Thalken, 2020; Reese, & Bhatia, 2018; Sarkar, 2019). 이러한 언어들은 NLP 작업을 수행하는 데 필요한 다양한 패키지와 라이브러리를 제공하며, 교육 분야에서의 텍스트와 음성 데이터 분석을 가능하게 해준다. 이를 통해 교육 자료의 개인화, 평가 자료의 생성, 교육적 대화의 이해, 텍스트 기반 분석 도구의 개발 등을 수행할 수 있다(Litman, 2016).

과학교육은 학생들에게 복잡한 개념과 이론을 전달하고, 학생들의 이해도와 학습 진행 상황을 적절하게 평가해야 한다. 교사 한 명이

* 교신저자 : 유선아 (sunaryu@knue.ac.kr)

<http://dx.doi.org/10.14697/jkase.2024.44.1.39>

많은 학생들의 이해도와 학습 진행을 파악하여 평가하는 것은 많은 시간과 노력이 요구되는 일인데, NLP 기법을 활용하면 평가를 매우 효율적으로 할 수 있어 과학교육 분야에서도 그 효용성에 대한 이해와 기대가 높아지고 있다. 예를 들어 학생들의 응답, 에세이, 그리고 교사의 피드백 등의 텍스트 데이터를 분석하는 데 큰 도움을 주고 있다(Wulff *et al.*, 2022). NLP 기법은 학생들의 열역학 관련 에세이를 분석하거나, 과학적 논증의 자동 점수 매기기 도구를 개발하는 데 활용된다(Donnelly, Vitale, & Linn, 2015; Lee, & Ryu, 2020). 또한 과학교육 분야의 연구나 보고서에 대한 자동화된 문헌 검토 및 분류를 수행하는 데에도 활용되고 있다(Odden, Marin, & Rudolph, 2021). 이러한 NLP 기법의 활용은 과학교육의 효과성과 효율성을 크게 높이며, 교사가 학생들의 학습 상황을 더욱 정확하게 파악하고, 개인화된 피드백과 지도를 제공할 수 있게 해준다(Wilson *et al.*, 2022). 또한 NLP 기법은 학생들의 과학적 사고 과정과 개념 이해도를 깊이 있게 분석하고(Ha *et al.*, 2011), 그 결과를 바탕으로 교육 전략을 개선하는 데 큰 역할을 하고 있다. NLP 기법을 과학교육에 활용하려면 평가되는 과학적 개념에 대한 깊은 이해가 필요하고, 이를 해결하기 위해서는 과학교육 전문가와 NLP 전문가 간의 협력이 중요하다(Zhai, He, & Krajcik, 2022).

과학교육 분야에서 NLP 기법을 적용하고자 노력하고 있지만, 구체적인 응용 및 활용 방법에 대한 체계적인 연구는 아직 충분히 이루어지지 않았다. 최근에는 우리나라도 NLP 기법과 관련된 연구가 점차 활발해지고 있지만, 연구 분야는 데이터를 쉽게 얻을 수 있는 문헌 검토 및 분류가 주를 이루고 있다. 이에 본 연구는 NLP 기법을 과학교육에 적용한 국내 및 해외 문헌들을 체계적으로 분석하여 NLP 기법의 연구 동향과 특징을 분석하고자 한다. 이를 위해 과학교육에서 NLP 기법의 주요 응용 분야를 분석하여 NLP가 어떻게 활용되는지 탐구한다. 그리고 교육자의 역할이 NLP 기법 도입에 따라 어떻게 변화하는지, 그 변화가 학습자에게 어떤 영향을 주는지 분석할 것이다. 또한 국내와 해외의 NLP 기법 활용 연구를 비교하고자 한다. 구체적인 연구 문제는 다음과 같다.

1. 과학교육 분야에서 NLP 기법의 주요 응용 분야는 무엇이며, 이러한 응용 분야에서 NLP는 어떻게 활용되고 있는가?
2. 과학교육 분야에서 NLP 기법의 도입이 교사의 역할에 어떠한 변화를 불러왔는가?
3. 국내와 해외의 과학교육 분야에서의 NLP 기법에 대한 연구는 어떠한 차이를 가지는가?

II. 이론적 배경

1. 자연어 처리(Natural Language Processing)

NLP는 컴퓨터 과학과 인공 지능의 분야에서 컴퓨터와 인간 언어 간의 상호 작용을 연구하는 핵심 영역이다. 디지털 라이브러리와 월드 와이드 웹의 급속한 성장에 따라, NLP의 중요성은 최근 몇 년 동안 크게 증가하였다(Chowdhary, & Chowdhary, 2020).

NLP는 텍스트 분류, 감정 분석, 기계 번역, 음성 인식 등과 같은 다양한 작업을 위해 컴퓨터가 인간의 언어를 이해, 해석 및 생성할

수 있도록 도와주는 알고리즘과 계산 모델을 개발한다. 20세기 중반부터 연구자들은 자연어 처리와 분석을 위한 컴퓨터의 활용 가능성을 탐구하기 시작했다. 초기에는 언어를 분석하고 생성하기 위한 규칙 기반 시스템이 주로 사용되었다. 1980년대와 1990년대에 들어서면서, 연구자들은 NLP 작업에 통계 모델과 기계 학습 알고리즘의 활용을 탐구하기 시작했다. 이러한 접근법은 대규모 언어 데이터셋에서 패턴과 관계를 학습하고, 그 학습을 기반으로 예측하는 데 중요하게 사용되었다. 최근에는 NLP 분야에서 딥 러닝 기술, 특히 신경망이 주목받고 있다. 이 기술은 언어 데이터를 처리하고 분석하기 위해 복잡한 계층적인 인공 뉴런 네트워크를 활용한다(Chowdhary, & Chowdhary, 2020).

2. 기계 학습(Machine Learning) 및 기계 학습 알고리즘(Machine Learning Algorithm)

기계 학습은 인공지능의 한 분야로 데이터를 통해 판단을 내릴 수 있는 알고리즘을 개발하는 학문이다. 이는 복잡한 패턴을 인식하여 예측을 수행하고 미래 상황에 대한 의사결정을 도와준다(Park *et al.*, 2015). 주요 과정에는 데이터 수집, 전처리, 모델 훈련, 평가, 그리고 개선이 포함된다. 특히 데이터 전처리 단계에서는 수집된 데이터를 학습할 수 있는 형태로 변환한다. 전처리된 데이터는 이후 훈련 데이터와 테스트 데이터로 나뉘어 모델 구축과 성능 평가에 사용된다(Yoo, 2019).

기계 학습은 학습 방식에 따라 지도 학습, 비지도 학습, 그리고 강화 학습으로 분류된다. 지도 학습은 입력 변수와 출력 변수 사이의 관계를 최적화하는 학습 방식으로 분류나 회귀 분석에 주로 사용된다. 반면, 비지도 학습은 올바른 정답이 없는 데이터를 기반으로 패턴을 발견하거나 데이터를 군집화하는 데 활용된다. 강화 학습은 보상을 최대화하는 방향으로 학습이 진행된다(Lantz, 2019).

기계 학습의 핵심은 데이터를 기반으로 새로운 예측이나 관계를 도출하는 것이다. 그러나 연구자가 데이터에 대한 깊은 지식이 없다면 모델의 결과를 올바르게 해석하기는 어려울 수 있다. 이러한 이유로 기계 학습을 활용하려는 연구자는 해당 분야의 전문 지식과 함께 기계 학습에 대한 깊은 이해가 필요하다. 특히 교육 분야에서는 교육학적 지식을 바탕으로 기계 학습 모델을 해석하고 적용하는 것이 중요하다고 강조된다(Yoo, 2019).

과학교육에 NLP 기법을 적용한 연구를 살펴보면 기계 학습 알고리즘으로 서포트 벡터 머신(SVM), 나이브 베이즈(NB), k-최근접 이웃 알고리즘(kNN) 등을 사용하였다. SVM은 분류와 회귀 작업에 적합한 기계 학습 알고리즘으로, 데이터를 균일하게 분할하는 초평면을 생성하는 것을 목표로 한다. 이 알고리즘은 다양한 학습 작업에 적용할 수 있으며, 특히 패턴 인식 분야에서 뛰어난 성과를 보여준다. NLP에서는 텍스트 데이터를 전처리하고, 단어 임베딩과 같은 기법으로 수치 특징으로 변환한 후 SVM 모델을 훈련해 새로운 텍스트 데이터에 대한 예측을 수행한다(Lantz, 2019). NB는 베이즈 정리를 기반으로 한 확률적 분류 알고리즘이다. 이 알고리즘은 모든 특징이 독립적이라는 단순화된 가정을 하지만, 텍스트 분류와 같은 다양한 작업에서 놀라게도 효과적이다. 훈련 과정에서는 레이블이 지정된 데이터셋을 사용하여 각 클래스에 대한 특징의 확률을 계산하며 이를 바탕으

로 새로운 데이터의 클래스를 예측한다. 특히 텍스트 데이터를 단어 수나 TF-IDF 점수와 같은 수치 특징으로 변환하여 사용하며, 이 간단한 접근 방식이 큰 특징 집합에서 높은 정확도를 달성하는 데 효과적이다(Lantz, 2019). kNN 알고리즘은 분류와 회귀 작업에 적합한 기계 학습 방법으로 유사한 예시들이 같은 클래스에 속한다는 원칙에 기반한다. 훈련 과정에서는 데이터를 저장하고 분류 시에는 새로운 데이터와 가장 가까운 k개의 이웃을 찾아 가장 흔한 클래스 레이블을 할당한다. 주로 이미지 인식 및 텍스트 분류와 같은 작업에 활용되며 간단하면서도 복잡한 작업에 효과적이다. 그러나 분류 속도가 느리고 많은 메모리를 요구하는 등의 단점도 가지고 있다(Lantz, 2019).

3. 딥 러닝(Deep Learning) 및 신경망 모델(Neural Network Model)

딥 러닝은 데이터의 표현을 학습하기 위해 여러 계층을 가진 인공 신경망을 사용하는 방법들의 집합이다. 딥 러닝의 핵심 아이디어는 입력 데이터로부터 특징의 계층 구조를 학습하는 것이며, 신경망의 각 계층은 데이터의 점점 더 복잡한 표현을 학습한다. 이를 통해 딥 러닝 모델은 스스로 해당 작업과 관련된 특징을 자동으로 학습할 수 있다(Deng, & Yu, 2014).

신경망 모델은 인간 두뇌의 구조와 기능에 영감을 받은 딥 러닝 모델의 한 유형이다. 신경망은 정보를 처리하는 연결된 노드나 뉴런의 계층으로 구성된다. 각 뉴런은 다른 뉴런으로부터 입력을 받아 비선형 활성화 함수를 입력에 적용하고, 다음 계층의 다른 뉴런에 전달되는 출력을 생성한다. 뉴런 간의 연결의 가중치와 편향을 조정함으로써 신경망은 복잡한 함수를 근사하고 새로운 데이터에 대한 예측을 학습할 수 있다(Deng, & Yu, 2014).

딥 러닝과 신경망 모델은 이미지 및 음성 인식, 자연어 처리, 자율 주행 등 다양한 응용 분야에서 큰 가능성을 보여주었다. 그러나, 이들은 훈련을 위해 대량의 데이터와 계산 자원이 필요하며 해석하고 디버깅(컴퓨터 프로그래밍에서 발생하는 오류나 결함을 찾아내고 수정하는 과정)하기 어려울 수 있다(Deng, & Yu, 2014).

과학교육에 NLP 기법을 적용한 연구를 살펴보면 LSTM 네트워크, Doc2Vec, BERT 모델 등의 신경망 모델을 사용하였다. LSTM 네트워크는 전통적인 RNN의 소실 기울기 문제를 해결하기 위해 개발되었으며, 이를 위해 장기 정보를 저장할 수 있는 메모리 셀과 정보를 제어하는 여러 게이트를 도입했다. 이 게이트들은 새로운 정보의 흐름, 정보의 저장 및 삭제, 그리고 네트워크의 출력을 제어한다. 훈련 과정에서는 입력 데이터를 기반으로 이 게이트와 메모리 셀의 가중치를 조절하며, 이를 통해 네트워크는 장기적인 데이터 종속성을 학습하게 된다. 이러한 특징으로 인해 LSTM은 언어 모델링, 기계 번역, 감정 분석 등의 시퀀스 데이터 분석에 매우 적합한 도구로 간주된다(Isotani *et al.*, 2019).

Doc2Vec은 Word2Vec의 발전된 형태로 말뭉치 내의 문서들을 벡터로 표현하게 학습한다. 이를 통해 유사한 내용의 문서들이 비슷한 벡터 표현을 가지게 되며, 이는 문서 분류, 클러스터링, 검색 등의 작업에 유용하게 사용된다. 이 모델은 크게 PV-DM과 PV-DBOW 두 가지 방식으로 변형될 수 있다. PV-DM은 단어의 주변 맥락과 문서의 벡터를 활용해 특정 단어를 예측하는 반면, PV-DBOW는 문

서의 벡터만을 이용해 예측을 수행한다. 각 방식은 문서 임베딩(텍스트 데이터를 컴퓨터가 이해하고 처리할 수 있는 숫자 형태의 벡터로 변환하는 과정)을 효과적으로 생성하지만 특정 작업이나 데이터셋에 따라 더 적합한 방식이 선택될 수 있다(Deng, & Yu, 2014).

BERT는 Google AI Language에서 개발한 사전 훈련된 언어 모델로 대량의 텍스트 데이터에서 비지도 학습을 통해 훈련되었다. 이 모델은 문장 내 단어 간의 문맥 관계를 학습하기 위해 설계되었으며, 다양한 자연어 처리 작업에 세부 조정될 수 있다. BERT는 텍스트 분류, 질문 응답, 명명된 엔터티 인식(텍스트에서 사람, 장소, 조직, 날짜 등과 같은 특정 정보를 식별하고 분류하는 과정) 등의 작업에서 최첨단 결과를 보여주었으며 학계와 산업 모두에서 널리 사용되고 있다(Wulff *et al.*, 2023).

4. NLP 기법

과학교육에 NLP 기법을 적용한 연구를 살펴보면 여러 기법이 도입되었는데, 그중에는 잠재 의미 분석(Latent Semantic Analysis, LSA), 잠재 디리클레 할당(Latent Dirichlet Allocation, LDA), n-gram, Escrito 프레임워크, Python SpaCy, SIDE 등이 있다.

LSA는 단어와 문서의 의미를 고차원 벡터로 분석하는 NLP 기법으로, 이 벡터의 차원을 줄이기 위해 특잇값 분해(SVD)를 활용한다. 이 과정을 통해 문서나 단어 간의 의미적 관계를 파악할 수 있다(Liu *et al.*, 2014).

LDA는 데이터의 패턴을 찾아내기 위한 비지도 학습 방식으로, 레이블이나 미리 정의된 결과 없이 작동한다. 사용자는 알고리즘을 실행하여 결과로 나타난 트렌드를 해석한다(Odden, Marin, & Rudolph, 2021).

n-gram은 연속된 단어들을 통해 텍스트 내 다음 단어를 예측하는 언어 모델이다. n의 크기에 따라 예측의 정확도는 높아지나, 훈련 데이터의 한계로 신뢰도는 감소할 수 있다(Brown *et al.*, 1992). 이 기법은 자연어 처리뿐만 아니라 DNA 서열 분석 등 다양한 분야에서 패턴 분석에 사용된다(Choi, Song, & Nam, 2010).

Liu *et al.*(2014) 연구에서는 Escrito 프레임워크를 소개하는데, 이는 학생의 글쓰기를 분석하고 피드백을 제공하는 NLP 도구다. 이 도구는 학생의 글에서 주요 아이디어를 식별하며, 글쓰기 기술 향상을 위한 피드백을 제공한다.

Python SpaCy는 텍스트 데이터로부터 언어적 특성을 추출하기 위한 NLP 라이브러리로 널리 알려져 있다. 이 라이브러리는 명명된 엔터티 인식, 품사 태깅, 의존성 구문 분석 등의 NLP 작업에 주로 사용된다. 과학교육 분야에서는 학생들의 텍스트 응답을 분석하여 그들의 과학적 개념에 대한 이해도를 평가하는 데 활용되었다(Sung *et al.*, 2021).

마지막으로, SIDE(Summarization Integrated Development Environment)는 카네기 멜런 대학에서 개발된 요약 통합 개발 환경이다. 이 시스템은 사용자 중심으로 개인화된 요약을 생성하며, 원본 문서에서 중요한 부분을 선택할 수 있다. 또한 GUI의 플러그인 관리자를 통해 사용자 정의 모듈을 추가하여 SIDE의 기능을 확장할 수 있다(Kang *et al.*, 2008).

5. 자동 평가 시스템

과학교육에 적용한 자동 평가 시스템을 살펴보면 HASbot, c-rater 등이 있다.

HASbot은 학생들의 논증을 기계 학습 기반의 NLP로 분석하고, 그 성과에 따른 실시간 피드백을 제공하는 시스템이다. 이 플랫폼은 5초 이내에 학생, 교사, 평가 도구, 그리고 교육용 서버와의 연결을 가능하게 하는 소프트웨어 아키텍처에 기반하고 있다. HASbot의 평가 모델은 교육 테스트 서비스에 의해 개발된 c-rater-ML이라는 NLP 엔진을 활용하여 구축되었다(Lee *et al.*, 2019).

c-rater™는 학생들의 내용 기반 짧은 답변을 즉시 평가하기 위해 자연어 처리를 사용하는 자동 평가 엔진이다. 이는 학생들의 개방형 질문에 대한 응답을 평가하기 위해 다양한 교육 환경에서 사용되었다. c-rater™는 단어 순서, 문자 순서, 의미 및 구문 종속성과 같은 다양한 언어적 특징을 기반으로 학생 응답의 질을 식별하도록 설계되었다(Zhu *et al.*, 2017). 반면, c-rater-ML은 짧은 구성 응답에 대한 평가를 위해 설계된 자동 평가 엔진이다. 이는 지도 학습 및 자동 모델 구축 과정을 사용하여 항목에 대한 평가 모델을 생성한다. c-rater-ML은 이전에 복잡한 구성 응답 과학 항목을 평가하는 데 사용되었다. 각 논증 블록에 대한 설명은 0에서 6까지의 7점 척도로 평가되며, 과학적 관련성, 상세 설명 및 추론의 다양한 수준을 나타낸다(Zhu *et al.*, 2017).

III. 연구 방법

본 연구는 국내 및 해외에서 진행되고 있는 NLP 기법의 과학교육 적용 연구의 최근 추세를 파악하기 위해 관련 연구 문헌을 수집하고 분석하였다. 이를 위해 문헌 조사와 내용 분석 방법론을 채택하였다. 연구 대상 논문의 검색은 학술 정보 데이터베이스인 구글학술검색, RISS, 그리고 국회 전자도서관에서 진행하였다. 검색어로는 ‘자연어 처리 기법+과학교육’, ‘Natural Language Processing+Science Educaiton’, ‘NLP+Science Education’, ‘텍스트 마이닝+과학교육’으로 하였다.

본 연구에서 논문 선정의 기준은 다음과 같다. (1) 국내 및 해외 학술지에서 발표된 과학교육 분야에서 NLP 기법을 활용한 연구, (2) 초·중·고등학생, 예비 교사 및 현직 교사를 대상으로 한 NLP 기법 적용 연구이다. 그러나 학술 대회에서 발표된 논문이나 과학 외의 다른 과목을 대상으로 한 연구는 연구 대상에서 제외하였다. 이러한 기준에 따라 선정된 논문은 2011년부터 2023년 9월까지의 총 37편이다(Table 1).

Table 1에 제시된 바와 같이 2011년부터 NLP 기법에 관한 연구가 지속해서 발표되어 왔으며, 최근 4년 동안 이 분야의 연구 활동이 특히 활발해진 것을 확인할 수 있다. Table 2에 따르면 NLP 기법을 활용한 과학교육 연구 논문 중 국내에서는 한국과학교육학회지에서, 해외에서는 Journal of Science Education and Technology에서 가장 많은 논문이 게재되었다.

과학교육 분야에서 NLP 기법의 활용 현황과 특징을 파악하기 위해 분석 대상 논문들을 ‘과학교육에서 NLP 기법의 주요 응용 분야’, ‘과학교육에서 NLP 기법을 이용할 때 교사의 역할’, ‘국내와 해외에서 이루어진 과학교육에서 NLP 기법을 이용한 연구 비교’ 측면에서 체계적으로 분석하였다. 먼저 과학교육에서 NLP 기법의 주요 응용 분야에 대한 분석은 분석 대상 논문들을 꼼꼼히 읽어보면서 bottom-up 방식으로 주요 카테고리를 생성하였고, 생성된 카테고리를 다시 리뷰하여 정밀화하였다. 분석의 카테고리로는 형성평가, 자동 채점, 문헌 검토 및 분류, 패턴 추출 등이 도출되었으며, 해당 카테고리를 기반으로 분석 대상 논문들의 주요 연구 내용을 엑셀 파일에 체계적으로 분류하였다. 과학교육에서 NLP 기법을 이용할 때 교사의 역할에 대한 분석은 분석 대상 논문들에서 교육에 NLP 기법을 적용하여 어떤 변화가 있었는지, 그 과정에서 교사는 어떤 역할을 해야 하는지에 대해 논의된 내용을 중심으로 진행하였다. 분석 대상 연구물을 반복적으로 읽으면서 교사의 역할과 관련된 부분을 추출하여 기록한 뒤, 이를 범주화하는 작업을 반복했다. 이를 통해 NLP 기법을 활용할 때 필요한 교사의 역할 기대를 포착하고자 하였다. 국내와 해외에서 이루어진 과학교육에서 NLP 기법을 이용한 연구 비교에 대한 분석은 형성평가, 자동 채점, 문헌 검토 및 분류, 패턴 추출 카테고리

Table 1. The status of domestic and international science education research related to natural language processing by year

년도	국내 학술지 논문 수	해외 학술지 논문 수	계
2011	0	1	1
2012	0	1	1
2013	0	1	1
2014	0	2	2
2015	0	1	1
2016	0	3	3
2017	0	1	1
2018	2	1	3
2019	0	1	1
2020	2	2	4
2021	3	4	7
2022	5	4	9
2023	1	2	3
계	13	24	37

Table 2. The status of domestic and international science education research related to natural language processing by source.

분류	수록지명	논문 수
국내 학술지 논문	과학교육연구지	1
	과학영재교육	1
	대한지구과학교육학회지	1
	영재교육연구	1
	초등과학교육	2
	한국과학교육학회지	5
	현장과학교육	1
해외 학술지 논문	The SNU Journal of Education Research	1
	CBE—Life Sciences Education	1
	EDUCATIONAL ASSESSMENT	1
	Educational Measurement: Issues and Practice	1
	Frontiers in Education	1
	International Journal of Artificial Intelligence in Education	1
	INTERNATIONAL JOURNAL OF SCIENCE EDUCATION	1
	Journal of Computer Assisted Learning	1
	JOURNAL OF RESEARCH IN SCIENCE TEACHING	1
	Journal of Research in Science Teaching	1
	Journal of Science Education and Technology	9
	Journal of the Learning Sciences	1
	PHYSICAL REVIEW PHYSICS EDUCATION RESEARCH	3
	Science Education	2
	계	37

고리에 대한 국내 논문과 해외 논문의 수를 양적으로 비교하였고, 국내와 해외의 연구 동향을 비교하여 서술하였다. 모든 과정에서는 과학교육 및 학습분석학 전문가 1인과 과학교육 전공 대학원생 1인이 함께 참여하는 정기적인 세미나를 통해 분석 과정과 분석 결과를 함께 검토하여 누락되거나 잘못된 내용들을 수정하였다. 또한 교차 검토를 통해 합의 과정을 거쳐 분석의 타당성을 확보하였다.

IV. 연구 결과 및 논의

1. 과학교육에서 NLP 기법의 주요 응용 분야

가. 형성평가

최근에는 NLP 기법이 과학교육의 다양한 영역에서 활용되고 있다. 이 연구에서는 국내 및 해외 학술지 37편을 분석하여 과학교육 분야에서 NLP의 주요 응용 분야를 탐색했다. 분석한 결과 NLP 기법이 과학교육에 활용되는 분야는 형성평가, 자동 채점, 문헌 검토 및 분류, 패턴 추출로 분류할 수 있다. 과학교육에 형성평가로서 NLP 기법을 활용한 연구를 분석한 결과는 Table 3과 같다.

분석한 논문들은 과학교육에서 NLP 기법을 형성평가의 도구로 활용할 수 있음을 보여준다. 이 연구들은 NLP를 활용하여 데이터를 분석하고, 이를 학생이나 수업 설계에 도움을 주는 것을 목표로 하였다. 과학교육 형성평가에 NLP를 활용한 문헌들에서 NLP 기법은 서로 다른 집단의 지식수준 차이를 구별하여 학생들의 수준에 맞는 수업 프로그램을 설계하는 데 도움을 준다(Wulff et al., 2023). 또한

NLP 기법은 학생들의 개방형 질문에 대한 답변을 분석하여 학생들의 생각과 학습 과정, 오개념에 대한 정보를 제공하여 교사들이 더욱 맞춤화되고 효과적인 피드백을 제공할 수 있다(Rosenberg, & Krist, 2020; Wilson et al., 2022). 그리고 NLP 기법을 적용한 과학 질문 답변 챗봇을 개발하여 학생들의 학습을 지원하는 도구로 활용할 수 있었다. 다만 이러한 챗봇은 사용한 NLP 기법에 따라 연속적인 대화가 불가능하거나 일상생활의 모든 대화를 반영하는 것에 어려움이 있다 (Min, & Yoo, 2022).

하지만 NLP가 과학 교육의 만병통치약이 아니며 교육자들이 NLP를 사용할 때 주의해야 한다고 경고한다(Wilson et al., 2022). 구체적으로 NLP는 아직 학생들의 복잡한 사고를 파악하는 데 어려움이 있고, 데이터의 편향이 존재할 수 있기에 교사가 제공하는 수준의 개인화된 피드백을 제공하기에는 어렵다. 그리고 NLP를 이용한 피드백 시스템을 개발할 때는 NLP, 기계 학습, 과학 지식 등 여러 구성 요소의 통합을 요구하므로 어려움이 많다(Zhu et al., 2017).

결론적으로 과학교육에서 NLP 기법의 활용은 형성평가의 효과를 극대화하며, 학생들의 학습 과정과 이해도를 깊이 있게 분석할 수 있지만 문제점 또한 존재하기에 NLP는 교사를 대체하는 것이 아닌 지원하는 도구로 사용해야 한다.

나. 자동 채점

과학교육에서 학생들의 과학적 지식과 논증 능력을 평가하는 것은 중요한 부분을 차지한다(Gombert et al., 2023). 최근에는 NLP 기법이 이러한 평가 과정에서 큰 도움을 주고 있으며, 특히 자동 채점 분야에서

Table 3. Research on utilizing NLP in science education for formative assessment

연번	저자(연도)	제목	NLP를 형성평가에 응용한 방법
1	Wulff <i>et al.</i> (2023)	Enhancing writing analytics in science education research with machine learning and natural language processing-Formative assessment of science and non-science preservice teachers' written reflections	물리학 및 비물리학 예비 교사들의 가상의 수업 상황에 대한 진술을 NLP 기법과 기계 학습을 활용하여 주제와 대표 지식을 식별함. 이를 통해 NLP 기법과 기계 학습을 사용하면 물리 및 비물리 예비 교사들 사이의 진술문에 대한 분석적, 형성적 평가 도구로 사용될 수 있음을 보여줌.
2	Wilson <i>et al.</i> (2022)	Classification of open-ended responses to a research-based assessment using natural language processing	예비 교사들에게 제공한 물리 문제 해결에 대한 개방형 응답을 분석하고 분류하기 위해 NLP 기법을 사용하며, 이를 통해 오개념을 파악하고 이해도를 높이기 위한 피드백을 제공함. BERT를 사용하여 힘, 가속도, 속도와 같은 물리 개념과 관련된 특정 키워드와 구문의 존재에 따라 응답을 다양한 카테고리 분류함.
3	Wulff <i>et al.</i> (2021)	Computer-based classification of preservice physics teachers' written reflections	예비 물리 교사들의 수업 방법에 대한 진술문 분석하기 위해 품사 태깅, 명명된 개체 인식, 감정 분석과 같은 NLP 기법을 활용함. 진술문을 수업 방법 설명, 수업 방법 분석, 수업 방법 평가, 미래의 수업 계획으로 범주화하고 분석하여 예비 교사들의 전문성 성장을 위한 피드백을 제공함.
4	Lee <i>et al.</i> (2021)	Machine learning-enabled automated feedback: Supporting students' revision of scientific arguments based on data drawn from simulation	6-12학년 학생들에게 과학적 개념을 가르치기 위해 사용한 시뮬레이션에서 얻은 데이터를 NLP 기법으로 분석함. NLP 기법을 통해 학생들의 시뮬레이션 상호작용과 눈중 점수 간의 관계, 그리고 과학 교실에서 적시에 지원을 제공하는 방법을 탐구함.
5	Rosenberg, & Krist(2020)	Combining machine learning and qualitative methods to elaborate students' ideas about the generality of their model-based explanations	중학생들의 화학 평가 문항에 대한 응답에서 인식론적 신념을 확인하고 개념 형성 과정에 대한 의미 있는 참여를 위한 구조도를 개발하는 데 NLP 기법을 사용함. NLP 기법과 기계 학습을 결합하여 학생들의 응답을 분석하고 구조도를 개발하여 학생들의 사고와 학습 과정을 더 잘 이해하고 학습을 지원하기 위해 더 목표 지향적인 피드백을 제공할 수 있음.
6	Lee <i>et al.</i> (2019)	Automated text scoring and real-time adjustable feedback: Supporting revision of scientific arguments involving uncertainty	HASbot이라는 NLP 기반의 형성 피드백 시스템을 소개함. 이 시스템은 중·고등학교 학생들의 과학적 논증을 분석하여 피드백을 제공하며, 이를 통해 학생들이 자신의 논증을 수정하고 글쓰기 기술을 향상할 수 있음을 보여줌.
7	Zhu <i>et al.</i> (2017)	Investigating the impact of automated feedback on students' scientific argumentation	NLP 기법을 사용하여 중학교 학생들의 기후 변화에 대한 과학적 논증을 분석함. c-rater™ 및 c-rater-ML 기술을 활용하여 학생들에게 즉각적인 채점 및 피드백을 제공했고, 이러한 자동 피드백이 학생들의 과학적 논증 학습에 긍정적인 영향을 미쳤음을 확인함.
8	Donnelly, Vitale, & Linn(2015)	Automated guidance for thermodynamics essays: Critiquing versus revisiting	과학교육에서 NLP 기법을 활용하여 중학생들의 열평형에 대한 에세이를 분석한 후 학생들에게 맞춤형 지도를 제공함. 이 피드백은 학생들의 설명을 개선하고 열과 온도에 대한 직관적인 아이디어를 과학적 지식과 언어로 일치시키는 데 도움을 주도록 설계되었음.
9	Min, & Yoo(2022)	Doc2Vec을 이용한 중학교 과학 질문-답변 챗봇 개발 및 학생 질의 분석	중학생들이 Doc2Vec 기반의 과학 질문-답변 챗봇을 활용한 기록을 분석함. 이러한 챗봇은 수업 시간 외에도 이용할 수 있어 원격수업이나 블렌디드 러닝 환경에서 학생들의 학습을 지원하는 도구로 활용될 수 있음.

Table 4. Research on utilizing NLP for automatic grading in science education

연번	저자(연도)	제목	NLP를 자동 채점에 활용한 방법
1	Gombert <i>et al.</i> (2023)	Coding energy knowledge in constructed responses with explainable NLP models	독일의 K-12 학생들이 작성한 에너지와 관련된 글쓰기를 NLP 기법으로 분석함. 이 연구에서는 학생들의 글쓰기에서 에너지 지식과 관련된 일곱 가지 핵심 개념의 존재 여부를 코드화하기 위해 채점표를 사용하여 평가함. 그리고 로지스틱 회귀, Escrito 프레임워크, LSTM 네트워크 등의 다양한 NLP 기법들이 학생들의 글쓰기에서 핵심 개념을 식별하고, 이를 얼마나 정확한 지식 네트워크로 표현하는지 평가함.
2	Zhai, He, & Krajcik(2022)	Applying machine learning to automatically assess scientific models	미국 중학생들이 광합성 현상을 설명하기 위해 작성한 모델과 모델에 대한 설명을 NLP 기법으로 분석함. 학생들이 개발한 모델에 대한 응답에서 특징을 추출하고, 해당 응답이 특정 개념을 포함 여부를 판별하는 알고리즘을 개발하고, 이를 통해 학생들에게 모델을 평가함. 연구 결과 NLP 기법을 활용한 알고리즘은 교사의 채점만큼 정확하게 학생 응답을 평가할 수 있음을 보여줌.
3	Sung <i>et al.</i> (2021)	How does augmented observation facilitate multimodal representational thinking? Applying deep learning to decode complex student construct	열에너지에 대한 학생들의 글쓰기에서 언어적 특징을 추출하여 이해도를 평가하기 위해 NLP 기법을 활용함. Python SpaCy를 사용하여 품사 및 명명된 엔터티 인식 특징을 구축하고, 이를 BERT 모델에 적용하여 학생들의 열에너지에 대한 이해도를 평가함.

연번	저자(연도)	제목	NLP를 자동 채점에 활용한 방법
4	Mao <i>et al.</i> (2018)	Validation of automated scoring for a formative assessment that employs scientific argumentation.	중학교 학생들의 기후 변화에 대한 형성평가의 답변을 c-rater-ML로 4가지 기준(주장 제기, 주장 설명, 주장의 불확실성 평가, 불확실성 평가의 근거에 대한 설명)에 따라 평가함. 연구 결과 c-rater-ML과 같은 자동 채점 도구가 학생들의 과학적 논증 기술을 정확하게 평가할 수 있음을 보여줌.
5	Liu <i>et al.</i> (2016)	Validation of automated scoring of science assessments	NLP 기법을 활용하여 고등학교 학생들의 물리학, 화학, 생물학에 대한 과학 탐구에 대한 응답을 c-rater-ML을 사용하여 자동 채점함. 연구 결과에 따르면 c-rater-ML은 교사의 채점과 높은 일치도를 보이며, 효율적으로 응답을 처리함.
6	Ha, & Nehm(2016)	The impact of misspelled words on automated computer scoring: A case study of scientific explanations	생물학 과정을 공부하는 대학생들을 대상으로 자연 선택, 유전학, 생태학과 관련된 개방형 과학 응답에 NLP 기법을 이용하여 자동 채점을 수행함. 응답에서 특징을 추출하고, 이를 바탕으로 SVM 및 NB 알고리즘과 같은 통계 모델을 구축하여 자동으로 채점함. 또한 잘못된 단어를 정확하게 식별하고 개방형 응답을 높은 정확도로 채점할 수 있는 자동 채점 시스템의 효과를 검증함.
7	Nakamura <i>et al.</i> (2016)	Automated analysis of short responses in an interactive synthetic tutoring system for introductory physics	NLP 기법을 활용하여 기초 물리학의 온라인 튜터링 시스템에서 개념적 질문에 대한 고등학생과 대학생들의 짧은 텍스트 응답을 자동 분류함. 이 연구에서 사용한 NLP 기법은 기계 학습 알고리즘을 훈련하는 방식의 자동 채점에 관한 것으로 c-rater-ML이라는 자동 채점 엔진을 중점적으로 다룸.
8	Liu <i>et al.</i> (2014)	Automated scoring of constructed-response science items: Prospects and obstacles	NLP 기법으로 초등학교와 중학생들의 과학 개념에 대한 응답을 분석하여 핵심 개념의 존재나 부재를 판단함. 이 연구에서 사용한 LSA, LDA, c-rater TM 과 같은 도구는 채점의 효율성과 일관성을 향상하지만, 교육 데이터의 질, 채점 루브릭의 복잡성, 학생 응답의 다양성 등의 요인에 따라 정확성이 영향을 받을 수 있다고 설명함.
9	Beggrow <i>et al.</i> (2014)	Assessing scientific practices using machine-learning methods: How closely do they match clinical interview performance?	NLP 기법을 활용하여 대학생들이 작성한 진화에 대한 설명을 자동으로 평가하는 기계 학습 도구의 효과성을 검토함. 이 도구는 학생들의 설명 텍스트를 분석하고 설명, 논증, 의사소통과 같은 과학적 실천의 수준을 평가하는 데 중점을 두고 있음.
10	Nehm, Ha, & Mayfield(2012)	Transforming biology assessment with machine learning: Automated scoring of written evolutionary explanations	NLP 기법을 활용하여 진화에 대한 대학생들의 설명을 자동으로 채점하는 방법을 소개함. SIDE 프로그램은 학생들의 설명을 분석하여 주요 개념과 관계를 식별하고, 이를 전문가가 생성한 요약과 비교함.
11	Ha <i>et al.</i> (2011)	Applying computerized-scoring models of written biological explanations across courses and colleges: Prospects and limitations	생물학을 공부하는 대학생들의 진화에 대한 응답 NLP 기법과 기계 학습을 결합한 SIDE 프로그램을 사용하여 자동으로 평가함. 연구 결과 이 도구는 전문가 채점과 높은 일치도를 보였으며, 교육자들이 학생의 작업을 효과적으로 평가하는 데 도움이 될 것으로 예상됨.
12	Lee, & Ryu(2021)	기계 학습을 활용한 논증 수준 자동 채점 및 논증 패턴 분석	고등학생들의 과학적 논증 답화를 자동 채점 모델을 활용하여 분석함. 논증 클러스터와 n-gram을 사용하여 과학적 논증의 실제 패턴을 자동 채점에 반영하여 성능을 개선하였고, 의사결정 나무와 랜덤 포레스트 기법을 통해 과학적 논증 수준에 따른 논증의 양상과 패턴을 분석함.
13	Lee, & Ryu(2020)	전문가의 형태소 분류를 활용한 과학 논증 자동 채점	분자 구조에 대한 고등학생의 과학 논증 수업 중의 발화를 NLP 기법과 기계 학습 알고리즘을 결합한 SIDE 프로그램을 사용하여 자동으로 평가함. 자동 평가는 교육자들에게 시간과 자원을 절약하며 학생 성과를 일관되게 평가할 수 있음을 보여줌.

그 중요성이 더욱 부각되고 있다. 과학교육에서 NLP 기법을 자동 채점하는데 활용한 연구들을 분석한 결과는 Table 4와 같다.

과학교육에서 학생들의 응답 평가는 교육의 질 향상에 중요한 역할을 한다. 전통적인 수기 채점은 시간이 많이 소요되며 객관성을 유지하기 어려운 문제점이 있다. 이를 해결하기 위해 NLP 기법과 기계 학습을 활용한 자동 채점 연구가 주목받고 있다. 분석 대상이 된 논문들은 크게 과학 개념 서술에 대한 자동 채점과 논증과 같은 개방형 응답에 대한 자동 채점 두 부분으로 나눌 수 있었다.

과학 개념 서술에 대한 자동 채점 연구 결과들을 살펴보면 과학 개념에 대한 핵심 개념의 유무를 판단하여 자동 채점을 수행한다. 그리고 이러한 자동 채점을 수행했을 때의 장점으로 Gombert *et al.*(2023)는 과학적 개념에 대한 응답을 NLP 기법으로 평가하여 시간과 자원을 절약하면서 학생들의 지식과 이해에 대한 귀중한 통찰을

제공한다고 설명한다. Zhai, He, & Krajcik(2022)는 NLP 기법을 이용하면 객관식에 의존했던 평가를 다양한 유형으로 대처하면서 복잡한 학생들의 사고를 이해할 수 있고, 기존 교실의 평가 관행을 개선할 수 있다고 강조한다. Nakamura *et al.*(2016)는 기초 물리학 온라인 튜터링 시스템에서 학생들의 짧은 응답을 NLP 기법을 활용하여 평가했을 때의 이점을 강조하면서 온라인 숙제 시스템이나 온라인 교육 시스템에 이를 적용하면 과학 교육의 수준과 효율성을 높일 수 있다고 설명한다. Liu *et al.*(2014)는 과학 교육에서 NLP 기법의 사용은 학생의 과학적 개념을 평가하는 데 있어 더 객관적이고 일관된 방법을 제공할 수 있으며, 영어 이외의 언어를 사용하는 다양한 문화적 배경을 가진 학생들을 평가하는 데 도움을 줄 수 있다고 주장하였다.

개방형 응답에 대한 자동 채점 연구 결과들을 살펴보면 개방형 응답에서는 교육자 또는 전문가가 미리 선정한 기준에 따라 학생들의

응답을 분류하여 평가한다. Beggrow *et al.*(2014)는 학생들의 진화에 대한 논증을 분석하여 설명, 논증, 의사소통과 같은 과학적 실천에 대한 평가를 제공했으며, 사람이 진행한 채점과 높은 상관관계를 보였다. Nehm, Ha, & Mayfield(2012)는 학생들의 진화에 대한 논증을 SIDE 프로그램으로 분석하여 주요 개념과 관계를 식별하고, 이를 전문가가 생성한 요약과 비교하였다.

종합적으로 NLP 기법은 과학 개념에 대한 서술이나 개방형 응답을 자동으로 평가할 수 있음을 보여준다. 다만, 이러한 평가가 성공적으로 이루어지기 위해서는 몇 가지의 개선이 필요하다. 기계 학습 모델을 훈련하기 위한 대량의 고품질 훈련 데이터가 필요하고 (Gombert *et al.*, 2023; Ha, & Nehm, 2016), 복잡한 과학적 개념과 추론을 정확하게 분석할 수 있는 더 정교한 NLP 기법이 개발되어야 한다(Ha, & Nehm, 2016). 이를 위해서는 과학교육 전문가와 NLP 전문가 간의 협력이 필요하다(Zhai, He, & Krajcik, 2022). 또한 NLP 기반 자동 채점을 지속해서 사용되기 위해 과학교육 전문가들은 NLP 기반 평가를 기존의 과학 교육과정과 교수법에 통합해야 할 것이다 (Mao *et al.*, 2018).

다. 문헌 검토 및 분류

과학교육 분야에서의 연구는 지속해서 발전하고 있으며, 이러한 연구의 추세와 흐름을 파악하는 것은 교육자들에게 중요한 정보를 제공한다. 최근에는 NLP 기법이 과학교육 연구에서 문헌 검토 및 분류에 활용되고 있으며, 이를 통해 연구의 주제나 범주를 효과적으로 분석할 수 있다. 과학교육에서 문헌 검토 및 분류에 NLP 기법을 활용한 연구를 분석한 결과는 Table 5와 같다.

NLP 기법이 과학교육 연구에서 문헌 검토 및 분류에 활용된 과학 교육 문헌들을 살펴보면, NLP 기법을 활용하여 연구 논문의 주제를 여러 범주로 분류하고, 연구의 흐름과 추세를 파악했다. 특히 연구 주제의 변화, 연구 방법론의 다양성, 그리고 연구의 주요 키워드와 테마를 중심으로 분석하였다(Chang, & Na, 2022; Kim, & Jhun, 2021; Odden, Marin, & Caballero 2020; Odden, Marin, & Rudolph, 2021).

이 연구들은 NLP 기법을 활용하면 대량의 문헌 데이터를 효과적으로 분석할 수 있으며, 연구의 주제나 범주를 정확하게 파악할 수 있다고 설명한다. 또한 시간에 따른 연구 주제의 변화나 연구의 흐름

Table 5. Research on utilizing NLP for literature review and classification in science education

연번	저자(연도)	제목	NLP를 문헌 검토 및 분류에 활용한 방법
1	Odden, Marin, & Rudolph(2021)	How has Science Education changed over the last 100 years? An analysis using natural language processing	LDA이라는 비지도 학습 기법을 사용하여 문헌 내 주제를 식별하고 분류함. 이를 통해 저널의 연구를 세 가지 주제 그룹으로 분류하고, 각 주제의 시간에 따른 변화를 정량적으로 분석함.
2	Odden, Marin, & Caballero(2020)	Thematic analysis of 18 years of physics education research conference proceedings using natural language processing	NLP 기법의 하나인 LDA 기법을 사용하여 문헌 검토 및 분류를 수행함. 18년 동안의 물리 교육 연구 회의(PERC) 발표문을 기반으로 주제를 식별하고 문서를 분류함.
3	Kim, & Oh(2023)	네트워크 텍스트 분석과 토픽모델링을 활용한 과학(영재)고 교육활동 내용 분석	네트워크 텍스트 분석과 LDA 기법을 활용하여 2021년 전국 과학고 및 영재학교의 교육활동을 분석함. 이 연구에서 네트워크 텍스트 분석을 통해 교육활동 주제어 간의 관계적 속성을 파악하고, 토픽 모델링을 통해 중등 과학영재 교육 활동의 중점사항을 6개의 주제로 분류함.
4	Shin(2022)	과학 교사의 과학·AI융합교육 실천 동기와 인식	15명 과학 교사의 과학 및 AI 융합 교육에 대한 실천 동기와 인식을 LDA 기법을 사용하여 분석함. 분석을 통해 실천 동기를 네 가지 주제로 분류했고, 데이터 분석과 과학실험에서 AI의 활용은 학생들에게 과학적 탐구 과정을 경험하게 하고, 이를 통해 과학교육의 효과를 향상할 수 있다고 인식함을 설명함.
5	Chang, & Na(2022)	잠재 디리클레 할당(LDA) 기반의 토픽 모델링 분석을 통한 ‘초등과학교육’ 학술지 연구논문의 주제 및 변화	1983년부터 2021년까지 초등과학교육 학술지 연구논문의 주제와 시간에 따른 변화를 LDA를 사용하여 분석함. 이 연구는 LDA 분석을 통해 추출된 14개 주제의 변화 추이를 알아보기 위해 히트맵으로 시각화했고, 이를 통해 초등과학교육 연구의 동향을 이해하고, 미래 연구의 방향 설정에 중요한 참고 자료가 될 수 있다고 논의함.
6	Chang, & Na(2022)	한국과학교육학회지는 44년간 어떤 주제로 어떻게 변화했는가? -잠재 디리클레 할당(LDA)을 활용한 토픽 모델링 분석-	한국과학교육학회가 44년 동안 발표한 논문들의 추세와 변화를 LDA를 사용하여 분석함. 이 연구는 44년간의 기간을 4년씩 11개 주기로 나누어 주기별 연구 주제들의 변화를 히트맵으로 시각화했고, 이를 통해 시간에 따라 변화하는 주제들을 파악하고, 각 주제에 대한 핵심 키워드 및 핵심 문서들과 연결하여 해석하고 논의함.
7	Jho, & Lee(2022)	머신러닝을 활용한 영재 학생들의 졸업논문 주제 경향 분석	2017년부터 2021년까지 S 영재학교 학생들의 졸업 연구 논문 초록을 BERT 기반의 NLP 기법으로 분석함. 분석을 통해 학생들의 연구는 7개의 주요 범주로 구분됨을 발견함.
8	Kim, & Jhun(2021)	토픽 모델링을 활용한 과학영재교육 연구 동향 분석	LDA를 활용하여 2015년부터 2019년까지의 과학영재 교육 관련 연구 동향을 분석함. 이 연구에서는 10개의 주제를 도출했고, 과거에는 특정 토픽의 비중이 상대적으로 높게 나타났으나, 최근으로 올수록 연구가 특정 주제에 치우치지 않고 고루 진행되고 있다는 것을 확인함.
9	Shin, Ha, & Lee(2018)	과학교육의 재미에 대한 재발견 -재미의 의미와 가치를 중심으로-	LDA 기법을 활용하여 국내에서 이루어진 재미에 대한 연구 동향을 거시적으로 확인함. 분석 결과 재미와 관련된 국내 학술지 논문들은 10개의 주제로 분류했고, 재미 관련 경험의 특성을 정리함. 그리고 과학교육에서 재미의 가치와 과학학습의 재미를 위한 과학교육의 지향점을 제시함.

을 파악함으로써 교육자들은 미래의 연구 방향을 설정하는 데 도움을 받을 수 있고 주장한다(Chang, & Na, 2022; Odden, Marin, & Caballero 2020; Odden, Marin, & Rudolph, 2021). NLP 기법을 활용하여 학생들의 연구 주제나 연구 방법을 분석함으로써 교육자들은 학생들에게 맞춤형 교육 방안을 제시할 수 있다(Jho, & Lee, 2022). 그리고 과학(영재)고의 교육과정 및 교육활동의 내용을 체계적으로 분석한 결과를 토대로 교육과정 및 교육활동을 개선하고, 이를 지원하는 정책을 수립할 수 있다(Kim, & Oh, 2023).

다만 과학교육 분야에서 NLP 기법의 사용은 비교적 새로운 분야이며, NLP 알고리즘의 정확성과 신뢰성에 대한 우려와 데이터 사용과 관련된 윤리적 문제점이 있다. 특히, 문헌 분석 및 분류에 주로 사용된 LDA는 과학교육 연구에서 복잡한 차이를 구별하는 데 어려움이 있기에 더욱 정교한 NLP 기법이 개발되어야 하고, 이를 위해서는 과학교육 전문가와 NLP 전문가 간의 협력이 필요하다(Chang, & Na, 2022; Odden, Marin, & Caballero 2020).

라. 패턴 추출

과학교육 연구에서는 학생들의 생각과 반응을 분석하기 위해 다양한 방법을 사용해 왔다. 최근에는 NLP 기법이 이러한 분석에 큰 도움을 주고 있다. 과학교육에서 학생들의 응답이나 설명들에서 나타나는 공통점을 찾거나 패턴을 추출하여 유형을 분류하는데, NLP 기법을

활용한 연구를 분석한 결과는 Table 6과 같다.

NLP를 패턴 추출에 활용한 연구를 살펴보면, Wulff *et al.*(2022)는 BERT라는 사전 훈련된 언어 모델을 사용하여 예비 물리 교사들의 글쓰기 반응을 분류하였다. 이를 통해 NLP 기법은 과학교육 연구에서 패턴 추출에 유용하게 사용될 수 있음을 보여준다. Lee *et al.*(2018)는 기계 학습 기술을 활용하여 학생들의 과학 논변 활동을 분석하였는데, 기계 학습을 활용한 분석은 교육 연구와 교실 현장에서의 실천에 크게 기여를 할 것으로 예상되었다. Sherin(2013)은 NLP 기법을 활용하여 예비 교사들의 글쓰기 반응에서 패턴을 추출하였고, NLP 기법은 과학교육 연구에서 복잡한 데이터 세트를 분석하는 데 통합적인 접근을 제공할 수 있음을 보여주었다.

과학교육에서 NLP 패턴 추출을 적용하면 학생들의 생각과 반응을 더 깊이 있게 이해하고, 정확하고 효율적으로 분석할 수 있다. 특히 대량의 텍스트 데이터를 빠르게 분석하여 학생들의 이해도나 오개념을 파악하는 데 큰 도움을 줄 수 있다(Wulff *et al.*, 2022). 이를 통해 교사는 교육 방법을 개선할 수 있고, 교육과정 개발이나 교과서 집필에도 긍정적인 영향을 미칠 수 있다. 또한 NLP 기법은 과학교육 연구에서 질적 및 양적 평가 방법 사이의 간극을 줄일 수 있어 연구의 통합성을 높일 수 있다(Wulff *et al.*, 2022).

다만 과학교육에서 NLP 및 기계학습을 사용할 때는 모델을 훈련하기 위한 대규모 훈련 데이터가 필요하다. 과학교육에서 데이터 수집은 많은 비용과 시간이 필요하기 때문에 훈련 데이터 크기를 선정하는데

Table 6. Research on utilizing NLP for pattern extraction in science education

연번	저자(연도)	제목	NLP를 패턴 추출에 활용한 방법
1	Wulff <i>et al.</i> (2022)	Bridging the gap between qualitative and quantitative assessment in science education research with machine learning-A case for pretrained language models-based clustering	BERT를 사용하여 물리학에서 표준화된 수업 상황에 대한 예비 교사들의 개방형 응답을 분류하고 패턴을 추출함. 이 방법이 질적 및 양적 평가 방법 사이의 격차를 줄이며 복잡한 데이터 세트 분석에 통합적인 접근법을 제공할 수 있다고 주장함.
2	Wulff <i>et al.</i> (2022)	Utilizing a pretrained language model (BERT) to classify preservice physics teachers' written reflections	과학교육에서 예비 교사들의 응답에 나타난 패턴을 추출하기 위해 NLP 기법을 활용함. NLP 기법을 통해 체계적으로 텍스트를 전처리하고, BERT를 사용하여 응답 내용의 패턴을 정확하게 분류함. 이를 통해 과학교육 전문가들은 예비 교사들의 학습과 발전에 대한 정보를 얻을 수 있음.
3	Sherin(2013)	A computational study of commonsense science: An exploration in the automated analysis of clinical interview data	중학생을 대상으로 한 계절 변화와 관련된 인터뷰 데이터를 분석하기 위해 NLP 기법과 벡터 공간 모델을 클러스터 분석과 결합하여 패턴 추출을 진행함. 이 방법을 통해 학생들의 답변에서 기본 지식 요소를 식별하였고, 이러한 패턴은 학생들의 과학 개념에 대한 추론을 이해하는 데 도움을 줌.
4	Oh, & Kang(2021)	온라인 과학 기사 텍스트 마이닝을 통해 분석한 에너지 용어 사용의 맥락	온라인 과학 기사에서 에너지 용어의 사용 패턴을 분석함. 언어 네트워크와 토픽 모델링을 통해 '기술', '연구', '개발' 등의 주요 용어와 그들 사이의 연결 패턴을 추출함. 이 연구를 통해 에너지라는 용어가 일상생활에서 어떻게 사용되는지 분석할 수 있었고, 이를 통해 에너지에 관한 과학 교육과정과 수업을 개선하는 데 사용될 수 있음을 보여줌.
5	Han, Kim, & Kim(2020)	텍스트마이닝을 활용한 대학생들의 외계 행성 개념 변화 연구	천문우주학과 대학생들의 외계 행성에 대한 개방형 질문에 대한 응답을 자연어 데이터를 사용하여 분석함. 연구는 학생들의 외계 행성에 대한 인식, 탐사 방법, 중요성, 그리고 교육적 필요성을 분석했고, 학생들의 외계 행성과 관련된 과학적 개념에 대한 오해와 이해의 격차를 식별할 수 있었음. 이를 통해 오개념과 격차를 해소하고 학생들의 과학적 개념 이해를 개선하는 교육 프로그램 및 자료 개발에 활용될 수 있다고 설명함.
6	Lee <i>et al.</i> (2018)	머신 러닝을 활용한 과학 논변 구성 요소 코딩 자동화 가능성 탐색 연구	과학교육에서 학생들의 논변 활동을 분석하기 위해 머신 러닝과 한국어 자연어 처리 기술을 활용함. 연구는 Toulmin의 논변 구성 요소를 기반으로 데이터를 분석하였고, kNN을 사용하여 가장 우수한 분석 성능을 보여줌. 그리고 과학 교육에서 NLP 기법의 활용은 교실에서의 논증 활동의 질을 향상하고 학생들의 비판적 사고와 추론 기술을 개선하는데 도움을 준다고 설명함.

신중하게 고려해야 한다(Wulff *et al.*, 2022). 그리고 과학교육 연구에서 NLP 기법을 효과적으로 사용하기 위해서는 학생들이 과학 개념을 논의할 때 사용하는 복잡하고 미묘한 언어를 정확하게 식별하고 분석할 수 있는 정교한 NLP 모델을 개발할 필요가 있다(Sherin, 2013).

2. 과학교육에서 NLP 기법을 활용할 때 교사의 역할

과학교육에서 NLP 기법의 도입은 Table 7과 같이 교사의 역할에 큰 변화를 주었다. 교사의 역할은 단순한 지식 전달자에서 학생들의 학습을 지원하고 촉진하는 지도자로 확장되었다(Lee *et al.*, 2019).

Table 7. Roles and requirements for teachers when utilizing NLP in science education

순번	제목	교사가 할 수 있는 역할	교사에게 더욱 요구되는 점
1	A computational study of commonsense science: An exploration in the automated analysis of clinical interview data	-	-
2	Applying computerized-scoring models of written biological explanations across courses and colleges: Prospects and limitations	교사들은 수업에서 다양한 질문에 대한 학생들의 응답을 수집함으로써 연구자들이 기계 학습 방법을 개선하고 생물학에서 핵심 개념을 전달하기 위해 과학적 언어가 어떻게 사용되는지에 대한 이해를 높이는 데 도움을 줄 수 있음.	교사들은 NLP 및 기계 학습 알고리즘 사용에 익숙해져서 이러한 도구들을 효과적으로 수업에 통합할 필요가 있음.
3	Applying machine learning to automatically assess scientific models	-	-
4	Assessing scientific practices using machine-learning methods: How closely do they match clinical interview performance?	교사들은 NLP 기반 도구들을 사용하여 학생들의 과학 실천의 질을 평가하고, 과학적 개념을 더 잘 이해하고 과학 실천에 더 효과적으로 참여하는 방법에 대해 피드백을 제공해야 함.	교사들은 기계 학습과 NLP 도구들의 결과를 해석하고 학생들의 과학 실천을 개선하는 방법에 대한 피드백을 제공할 수 있어야 함.
5	Automated analysis of short responses in an interactive synthetic tutoring system for introductory physics	교사들은 NLP 기반 도구를 사용하여 응답을 자동으로 채점하고, 즉각적인 피드백을 제공하여 가르치는 개념을 더 잘 이해하도록 지도해야 함. 또한 NLP 기반 도구를 사용하여 학생들의 응답에서 패턴을 식별함으로써 학생들이 직면하고 있는 오개념과 어려움을 더 잘 이해하고 그에 따라 교수법을 조정해야 함.	교사들은 NLP 기반 도구 사용법을 교육받는 동시에, 이 분야의 최신 발전에 발맞춰 전문성을 지속적으로 개발해야 함. 또한 교사들은 NLP 기반 도구의 개발과 개선에 더 많이 참여해야 하며, 도구의 효과성에 대한 피드백을 제공하고 개선이 필요한 영역을 식별하는 데 기여해야 함.
6	Automated guidance for thermodynamics essays: Critiquing versus revisiting	교사들은 NLP 프로그램이 제공하는 자동 가이드를 사용하여 학생 학습을 더 개인화되고 효율적인 방식으로 지원해야 함.	교사들은 NLP 기법이 학습을 지원하는 데 어떻게 사용할 수 있는지 알아야 함. 또한 NLP 프로그램이 제공하는 가이드가 학생들의 기존 지식과 일치하는지를 확인하고 NLP 기법의 사용이 교실에서 교사의 역할을 대체하지 않도록 해야 함.
7	Automated scoring of constructed-response science items: Prospects and obstacles	교사들은 자동 채점 도구가 제공하는 데이터를 활용해 학생들이 추가적인 지원이 필요한 부분을 파악하고, 그에 따라 교수 방법을 조정해야 함.	NLP 알고리즘이 더 정교하고 복잡해짐에 따라 교사들은 이 알고리즘을 이해하고 해석하여 학생 학습을 지원할 수 있어야 함. 또한 데이터 과학자 및 기타 전문가와 협력하여 과학 교육의 요구에 맞게 NLP 알고리즘을 개발하고 개선할 수 있어야 함.
8	Automated text scoring and real-time adjustable feedback: Supporting revision of scientific arguments involving uncertainty	교사들은 HASbot과 같은 NLP를 활용한 형성적 피드백 시스템을 사용하여 설명, 논증, 의사소통과 같은 과학 실천을 다양한 과목과 학년에서 촉진해야 함.	교사들은 NLP에 대한 역량과 인식, 그리고 그것이 학생 학습 성과에 미치는 영향을 이해해야 함. 또한 교사들은 HASbot과 같은 NLP를 활용한 형성적 피드백 시스템을 효과적으로 사용하는 방법과 교실 수업에 원활하게 통합하는 방법에 대해 교육받아야 함.
9	Bridging the gap between qualitative and quantitative assessment in science education research with machine learning—A case for pretrained language models-based clustering	교사들은 NLP 기법을 사용하여 과학 관련 질문에 대한 학생들의 응답을 분석하여 사고 패턴을 식별하고 과학 개념에 대한 이해를 향상하기 위한 피드백을 제공할 수 있음.	교사들은 교실에서 NLP 기법을 효과적으로 사용하는 방법을 배워야 함. 또한 교사들은 NLP 분석의 결과를 해석하고 이를 수업에 반영하는 방법을 알아야 함.
10	Classification of open-ended responses to a research-based assessment using natural language processing	교사들이 NLP를 사용하여 학생 학습과 오개념에 대한 이해를 얻고, 이해를 향상하기 위한 목표 지향적인 피드백을 제공하며, 대량의 비정형 데이터를 보다 효율적으로 분석할 수 있음.	교사들은 NLP와 기계학습 기술들을 효과적으로 분석하고 해석하기 위해 기술에 대한 깊은 이해가 필요함. 또한 교사들은 학생들의 반응을 정확히 포착하고 특정 학생 그룹에 편향되지 않도록 NLP 모델의 개발과 훈련에 더 많이 관여해야 함.
11	Coding energy knowledge in constructed responses with explainable NLP models	교사들은 NLP 기법을 사용하여 학생들의 응답을 분석하고 학생들이 추가 지원이 필요할 수 있는 영역을 식별해야 함.	교사들은 NLP 기법을 이용하여 교육 데이터를 수집하고, 기계 학습 모델을 훈련하고, 이러한 모델의 결과를 해석하는 방법을 이해해야 함.

순번	제목	교사가 할 수 있는 역할	교사에게 더욱 요구되는 점
12	Combining machine learning and qualitative methods to elaborate students' ideas about the generality of their model-based explanations	교사들은 NLP를 사용하여 학생들이 어려워하거나 오해할 수 있는 분야를 이해하고, 주제를 더 효과적으로 이해할 수 있도록 개인화된 지원을 제공할 수 있음.	교사들은 NLP 기법을 본인의 수업에 효과적으로 사용하기 위해 새로운 기술과 지식을 개발해야 함.
13	Computer-based classification of preservice physics teachers' written reflections	교사들은 NLP를 사용하여 글쓰기 보조 도구, 평가 시스템, 서면 반응의 구조와 내용에 대한 피드백을 제공할 수 있는 지능형 컴퓨터 기반 튜터를 개발할 수 있음.	교사들이 NLP에 대한 기본적인 지식을 이해하고 있어야 하며, 이러한 방법이 전문적 성장을 지원하는 데 어떻게 사용될 수 있는지 알아야 함. 또한 개별 학생들의 필요에 더 정확하고 구체적으로 대응하는 피드백을 제공하기 위해 NLP 기반 도구를 사용할 수 있어야 함.
14	Enhancing writing analytics in science education research with machine learning and natural language processing-Formative assessment of science and non-science preservice teachers' written reflections	-	-
15	How does augmented observation facilitate multimodal representational thinking? Applying deep learning to decode complex student construct	교사들은 NLP 기반 자동 평가 도구가 생성한 결과를 사용하여 학생의 표현적 사고 모드를 구별하고 적시에 피드백을 제공해야 함.	교사들은 NLP 기반 자동 평가 도구를 사용하기 필요한 기술과 지식을 갖추도록 지속적인 전문적 개발과 훈련이 필요함.
16	How has Science Education changed over the last 100 years? An analysis using natural language processing	-	-
17	Investigating the impact of automated feedback on students' scientific argumentation	교사들은 과학 교육에서 자동 채점 및 피드백 시스템의 사용하여 학생들의 학습을 지원하고 과학적 논증 기술의 발달을 도와줘야 함.	교사들은 과학에서 학생 학습을 지원하는 데 NLP 기법을 어떻게 사용할 수 있는지 더 깊은 이해가 필요함.
18	Machine learning-enabled automated feedback: Supporting students' revision of scientific arguments based on data drawn from simulation	교사들은 과학 수업에서 NLP를 사용하여 학생들에게 즉각적인 피드백을 제공할 수 있음.	교사들이 NLP 및 기계학습 알고리즘으로부터 제공되는 진단 정보를 효과적으로 해석하고 사용할 수 있어야 함.
19	The impact of misspelled words on automated computer scoring: A case study of scientific explanations	-	-
20	Thematic analysis of 18 years of physics education research conference proceedings using natural language processing	-	-
21	Transforming biology assessment with machine learning: Automated scoring of written evolutionary explanations	교사들은 NLP 기법과 기계 학습 알고리즘을 사용하여 학생들의 서면 설명의 텍스트를 분석하고 그들 사이의 주요 개념과 관계를 식별할 수 있어야 함. 이를 통해 교사들은 학생들의 설명에서 흔히 발생하는 오해와 오류를 파악할 수 있으며, 이러한 정보는 학생 학습을 개선하기 위한 교육 자료와 개입의 설계에 반영해야 함.	교사들은 NLP 기법과 기계 학습 도구들을 효과적으로 사용할 수 있도록 전문을 개발해야 함. 또한 교사들은 기존의 교수 방법에 NLP 기법과 기계 학습 알고리즘을 통합하여 학생 학습에 대한 영향을 극대화할 수 있어야 함.
22	Utilizing a pretrained language model (BERT) to classify preservice physics teachers' written reflections	교사들은 NLP 기법을 활용하여 학생들의 반응을 체계적이고 객관적으로 분석할 수 있으며, 학생들에 추가 지원이나 훈련이 필요한 영역을 식별하여 더 효과적인 교수 전략 개발해야 함.	교사들은 기계학습과 자연어 처리에 대한 기본적인 이해뿐만 아니라 관련 소프트웨어와 도구를 사용할 수 있는 능력을 갖추어야 함. 또한 NLP 분석 결과를 해석하고 이를 수업에 반영하는 방법과 NLP 기법을 사용할 때 윤리적 고려 사항을 인지해야 함.
23	Validation of automated scoring for a formative assessment that employs scientific argumentation	교사들은 자동 채점 도구를 활용하여 학생들에게 즉각적인 피드백을 제공하고, 학생들이 과학적 개념과 주제에 대한 이해를 개선해야 할 분야를 파악해야 함.	교사들은 NLP 기법과 자동 채점 도구를 교수 실습에 효과적으로 통합하기 위해 이 기술들에 더 익숙해져야 함. 또한 교사들은 자동 채점 도구 개발자들과 협력하여 이러한 도구가 그들의 교육과정 및 교육 목표와 일치하도록 해야 함.
24	Validation of automated scoring of science assessments	교사들은 자동 채점으로 생성된 피드백을 사용하여 학생들이 어려워하는 부분을 식별하고 그에 따라 교수 전략을 조정해야 함.	교사들은 NLP 도구를 자신의 수업에 통합하고 학생들에게 더 개인화된 피드백을 제공하는 데 사용할 수 있어야 하며, 교육과정의 특정 요구에 맞춤형된 NLP 도구를 만들기 위해 개발자들과 협력해야 함.
25	Doc2Vec을 이용한 중학교 과학 질문-답변 챗봇 개발 및 학생 질의 분석	교사들은 NLP 기법을 활용하여 교육용 챗봇을 개발하고, 학생들이 챗봇을 사용하면서 발생하는 질의응답 데이터를 분석할 수 있음. 그리고 교사는 챗봇을 통해 학생들의 질문에 적절하게 응답할 수 있는 데이터셋을 구축하고, 지속적으로 업데이트하여 챗봇의 성능을 향상시킬 수 있음.	교사들은 NLP 및 기계학습 기술에 대한 이해와 지식을 갖추어야 함. 교사들은 챗봇이 학생들의 필요에 부응하도록 챗봇의 기능을 지속적으로 개선하고 업데이트하는 역할을 수행해야 함.

순번	제목	교사가 할 수 있는 역할	교사에게 더욱 요구되는 점
26	과학 교사의 과학·AI융합교육 실천 동기와 인식	교사들은 토픽 모델링과 같은 NLP 기법을 사용하여 과학·AI 융합 교육의 동기와 인식을 분석하고 과학 교사들의 어려움과 요구 사항을 식별할 수 있음. 이는 AI를 과학 교육에 효과적으로 통합하는 교육과정과 교재 개발에 도움이 될 수 있음.	교사들은 NLP 기법을 사용하여 데이터를 분석하고 과학 교육 및 AI 융합 교육을 개선할 수 있는 영역을 식별할 수 있어야 함. 이를 위해서는 AI와 NLP의 최신 발전을 따라잡기 위한 지속적인 전문 개발과 훈련이 필요함.
27	과학교육의 재미에 대한 재발견 -재미의 의미와 가치 중심으로-	-	-
28	기계 학습을 활용한 논증 수준 자동 채점 및 논증 패턴 분석	교사들은 자동 채점 모델을 활용하여 학생들의 논증 수준을 신속하고 정확하게 평가할 수 있고, 이를 활용하여 개별 학생의 논증 능력에 맞춘 피드백을 제공할 수 있음. 그리고 학생들의 논증 패턴을 분석하고, 이를 교육과정에 적용할 수 있음.	교사들은 NLP와 자동 채점 기술에 대한 이해와 지식을 갖추어야 하며, 자동 채점 결과를 학생 지도에 효과적으로 활용할 수 있는 능력을 개발해야 함. 그리고 학생들의 논증 패턴 분석 결과를 교육과정과 수업 방법에 통합하는 방법을 배워야 함.
29	네트워크 텍스트 분석과 토픽모델링을 활용한 과학(영재)고 교육활동 내용분석	-	-
30	머신 러닝을 활용한 과학 논변 구성 요소 코딩 자동화 가능성 탐색 연구	교사들은 NLP 기법을 사용하여 학생들의 논증 활동을 분석함으로써 논증의 인지적 측면을 더 잘 이해하고 학생들의 논증 기술의 강점과 약점을 파악할 수 있고, 이를 이용하여 더 효과적인 교수 전략을 개발하고 학생들의 논증 기술을 향상하기 위한 목표를 가진 피드백을 제공해야 함.	교사들은 자연어 데이터를 분석하기 위해 NLP 도구와 자원을 사용하는 능력뿐만 아니라 이러한 분석 결과를 해석하고 교육 및 학습 성과를 개선하기 위해 적용하는 능력을 개발해야 함.
31	머신러닝을 활용한 영재 학생들의 졸업논문 주제 경향 분석	학생들의 연구 추세와 흐름을 분석하고, 이를 지도에 활용하여 교육 방향성을 제공할 수 있고, 학생들의 연구 주제를 범주화하여, 유사한 관심사를 가진 학생들 간의 협력과 교류를 촉진할 수 있음. 또한 졸업 연구와 관련된 데이터셋을 수집하여 학생들에게 더욱 나은 연구 가이드라인을 제시할 수 있음.	학생들의 연구 데이터를 정확하게 분석하고 해석하는 데 필요한 NLP 기법에 대한 이해가 요구되며, NLP 기반 분석 도구를 활용하여 학생들의 연구 프로젝트를 효율적으로 관리하고 지도하는 방법을 습득해야 함.
32	온라인 과학 기사 텍스트 마이닝을 통해 분석한 에너지 사용의 맥락	온라인 과학 뉴스 기사에서 자연어 데이터를 분석함으로써 교사들은 과학적 개념이 실제 세계에서 어떻게 사용되고 이해되는지에 대한 정보를 얻을 수 있음. 이는 학생들에게 관련성이 있고 의미 있는 과학적 개념을 가르치는 데 있어 더욱 흥미롭고 효과적인 접근법을 개발할 수 있음.	교사들은 자연어 데이터를 분석하기 위한 NLP 기법 및 기타 컴퓨터 방법론에 더욱 익숙해질 필요가 있음. 또한 컴퓨터 언어학 및 데이터 분석 전문가와 긴밀히 협력하여 자연어 데이터를 분석하기 위한 새로운 방법과 도구를 개발하고 분석에 사용할 수 있는 고품질 데이터 소스를 식별할 필요가 있음.
33	잠재 디리클레 할당(LDA) 기반의 토픽모델링 분석을 통한 '초등과학교육' 학술지 연구논문의 주제 및 변화	-	-
34	전문가의 형태소 분류를 활용한 과학 논증 자동 채점	교사들은 자동 채점 모델을 활용하여 학생들의 과학적 논증 수준을 효과적으로 평가하고 실시간 피드백을 제공할 수 있음. 그리고 자동 채점의 결과를 활용하여 논증 수업의 설계 및 피드백 모델을 개발할 수 있음.	교사들은 기계 학습 및 자동 채점 기술에 대한 충분한 이해를 갖추어야 함. 그리고 학생들의 논증 발화 데이터를 효과적으로 활용하고, 이를 교육과정에 통합하는 방법을 습득해야 함.
35	텍스트마이닝을 활용한 대학생들의 외계행성 개념 변화 연구	-	-
36	토픽 모델링을 활용한 과학영재교육 연구동향 분석	-	-
37	한국과학교육학회지는 44년간 어떤 주제로 어떻게 변화했는가? -잠재 디리클레 할당(LDA)을 활용한 토픽모델링 분석-	-	-

-: 해당 논문에서 내용을 구체적으로 제시하지 않음

NLP 기법을 수업에 도입하면 교사들은 학생들의 학습과 오개념에 대한 정보를 얻을 수 있고, 이를 해결하기 위한 개별 맞춤형 피드백을 제공할 수 있다(Nakamura *et al.*, 2016). 또한 NLP를 활용하여 형성적 피드백을 제공하여 학생들의 과학적 논증 기술을 향상하는 방법에 대한 지침을 제공하며(Lee *et al.*, 2021), 학생들의 응답을 분석한 후, 학생들에게 필요한 피드백을 식별할 수 있다(Zhai, He, & Krajcik, 2022). 그리고 교사들은 자동 채점 시스템의 결과를 활용하여 교수법

을 개선하는 방안을 모색할 수 있다(Ha, & Nehm, 2016).

NLP 기법의 도입은 또한 과제 및 활동 설계에도 큰 영향을 미쳤다. 교사들은 학습 목표와 일치하는 과제를 설계하며, 이를 통해 학생들의 학습 증거를 수집한다(Lee *et al.*, 2021). 더 나아가 교사들은 모델링 작업을 설계하고, 학생들의 응답을 수집하여 분석하는 역할을 수행한다(Zhai *et al.*, 2022). 교사들이 NLP를 과학교육에 효과적으로 활용하기 위해서는 NLP 기법의 장점과 한계를 깊이 이해하고, 학생

들의 개별적인 필요에 따라 적절한 지도를 제공해야 한다(Liu *et al.*, 2016).

교사들은 NLP 기반 자동 채점 도구의 개발 및 활용에서도 중요한 역할을 수행한다. Oh & Gang(2021)의 연구에서 교사들은 에너지 관련 기사에서 발견된 용어 사용 패턴을 바탕으로 에너지 관련 수업에서 학생들의 진로 지도 및 관련 분야의 연구 소개 기회를 제공할 수 있었다(Oh & Gang, 2021). 또한 교사들은 NLP 도구의 설계, 구현, 그리고 생성된 점수의 해석 및 사용에 대한 의견을 제공해야 하며(Mao *et al.*, 2018), NLP 기법을 활용하여 얻어낸 데이터를 분석하고 처리하는 능력을 갖추어야 한다.

NLP 기법의 사용으로 인한 교사들의 역할 변화는 학생들에게도 긍정적인 영향을 미친다. 학생들은 즉각적인 피드백을 통해 자신의 학습 성과를 개선할 수 있으며(Donnelly, Vitale, & Linn, 2015; Lee *et al.*, 2021; Liu *et al.*, 2016; Mao *et al.*, 2018; Zhai, He, & Krajcik, 2022), 자기주도 학습을 통해 학습의 주체성을 더욱 강화할 수 있게 되었다(Min, & Yoo, 2022; Zhai, He, & Krajcik, 2022). NLP 기법이 활용된 자동 채점을 통해 학생들은 마치 여러 명의 교사가 동시에 수업을 지원하는 것처럼 느낄 수 있고, 즉각적인 피드백을 받을 수 있는데, 이는 교사가 직접 지도하는 것보다 훨씬 효율적이다(Donnelly *et al.*, 2015).

또한 과학적 논증 및 글쓰기 기술, 과학적 소양과 개념에 대한 이해, 그리고 연구 능력 및 전략 등 다양한 영역에서의 향상을 경험하게 되었다(Lee, *et al.*, 2018; Lee *et al.*, 2019; Lee, & Ryu, 2021). Lee & Ryu(2021)의 연구에서 교사들은 학생들에게 HASbot의 사용 방법을 제시하여, 물 모듈을 구현하는 과정에서 학생들을 적극적으로 지원하고 격려했다. 이로 인해 학생들은 불확실성을 포함한 과학적 논증 상황에서 더욱 효과적인 피드백을 받게 되었고, 학생들의 논변 활동의 질을 향상시키는데 큰 도움을 주었다(Lee & Ryu, 2021).

이렇듯 NLP 기법의 도입은 과학교육의 질과 효과성을 크게 향상시키는 중요한 역할을 할 것으로 기대된다. 그러나 과학교육에 NLP를 성공적으로 도입하기 위해서는 교사들이 NLP와 기계학습을 효과적으로 분석하고 해석할 수 있는 능력을 갖추어야 한다(Donnelly, Vitale, & Linn, 2015; Ha *et al.*, 2011; Wulff *et al.*, 2021). 이를 위해서 교사들은 NLP와 기계학습과 같은 최신 기술을 활용하기 위해 지속적으로 전문성을 개발해야 한다(Nakamura *et al.*, 2016; Nehm, Ha, & Mayfield, 2012; Sung *et al.*, 2021). 그리고 교사들은 NLP 기법자들과

Table 8. Comparison of domestic and international research counts based on the purpose of using NLP in science education

NLP를 활용한 목적	국내 연구 수 (편)	해외 연구 수 (편)
형성평가	1	8
자동 채점	2	11
문헌 검토 및 분류	7	2
패턴 추출	3	3
계	13	24

협력하여 NLP 도구가 개선되도록 피드백을 제공해야 하며(Nakamura *et al.*, 2016), 과학 교육과정 및 교육 목표에 일치하도록 노력해야 한다(Liu *et al.*, 2014).

3. 국내외에서 이루어진 NLP 기법을 활용한 과학교육 연구 비교

과학교육에서 NLP 기법을 활용한 국내외 해외 연구를 NLP 기법을 활용한 목적에 따라 세부적으로 비교해 보면 Table 8과 같다.

2011년부터 2023년 9월까지 과학교육 분야에서 NLP 기법을 활용한 연구가 총 37건 발표되었는데, 이 중 국내 연구는 13건이고 해외 연구는 24건으로 확인되었다. 해외에서는 형성평가 8건, 자동 채점 11건, 문헌 검토 및 분류 2건, 패턴 추출 3건의 연구가 발표되었다. 해외에서는 형성평가와 자동 채점 분야에서 NLP 기법이 특히 활발하게 활용되고 있다.

한편, 해외 연구와 비교할 때 국내에서의 NLP 활용 연구는 초기에는 상대적으로 적었는데 최근에 활발하게 진행되고 있다. 다만, 주로 문헌 검토 및 분류에 활용되는 모습을 볼 수 있다.

해외에서 이루어진 연구 동향을 분석하면 Table 9와 같다. 연구 대상을 살펴보면 예비 교사, 초등학생, 중학생, 고등학생, 대학생 등 다양한 대상에서 NLP 관련 연구가 수행되었다. 이는 NLP 기법이 모든 대상에 걸쳐 유용하게 적용될 수 있음을 보여준다. 또한 연구에서는 물리학, 화학, 생명과학, 지구과학 등과 같이 과학의 여러 하위 분야에서 활용되고, 일부 연구에서는 1,000명 이상의 대규모 데이터를 다루고 있음을 보여준다. 이는 NLP가 대용량 데이터를 분석하는데 효과적인 도구임을 나타낸다.

Table 9. International research trends in NLP

순번	제목	연구 대상	분석 프로그램	과학 내 과목	연구 대상 수
1	A computational study of commonsense science: An exploration in the automated analysis of clinical interview data	중학생	-	지구과학	54명
2	Applying computerized-scoring models of written biological explanations across courses and colleges: Prospects and limitations	대학생	SIDE	생명과학	-
3	Applying machine learning to automatically assess scientific models	고등학생	로지스틱 회귀, 의사결정 나무, 랜덤 포레스트	생명과학	-
4	Assessing scientific practices using machine-learning methods: How closely do they match clinical interview performance?	대학생	NLP와 기계학습을 결합하여 자체적으로 개발한 자동 채점 모델	생명과학	104명

순번	제목	연구 대상	분석 프로그램	과학 내 과목	연구 대상 수
5	Automated analysis of short responses in an interactive synthetic tutoring system for introductory physics	고등학생, 대학생	LSA, 로지스틱 회귀, 서포트 벡터 머신, 의사결정 나무	물리학	고등학생: 41명 대학생: 161명
6	Automated guidance for thermodynamics essays: Critiquing versus revisiting	중학생	c-rater™	물리학	120명
7	Automated scoring of constructed-response science items: Prospects and obstacles	중학생	LSA, c-rate	물리학, 생명과학, 지구과학	-
8	Automated text scoring and real-time adjustable feedback: Supporting revision of scientific arguments involving uncertainty	중·고등학생	n-gram, LDA	지구과학	-
9	Bridging the gap between qualitative and quantitative assessment in science education research with machine learning-A case for pretrained language models-based clustering	예비 물리 교사	-	물리학	-
10	Classification of open-ended responses to a research-based assessment using natural language processing	예비 교사	토큰화, 스테밍, 불용어 제거, k-클러스터링, SVM	물리학	-
11	Coding energy knowledge in constructed responses with explainable NLP models	중·고등학생	BERT, LSTM 네트워크, 로지스틱 회귀, 의사결정 나무, 랜덤 포레스트, SVM, kNN	물리학	305명
12	Combining machine learning and qualitative methods to elaborate students' ideas about the generality of their model-based explanations	중학생	LDA	화학	173명
13	Computer-based classification of preservice physics teachers' written reflections	예비 교사	품사 태깅, 명명된 개체 인식, 감성 분석	물리학	81명
14	Enhancing writing analytics in science education research with machine learning and natural language processing-Formative assessment of science and non-science preservice teachers' written reflections	예비 교사	-	물리학	-
15	How does augmented observation facilitate multimodal representational thinking? Applying deep learning to decode complex student construct	-	Python SpaCy, LDA, SVM	물리학	-
16	How has Science Education changed over the last 100 years? An analysis using natural language processing	과학교육 문헌	LDA	-	약 5,577편
17	Investigating the impact of automated feedback on students' scientific argumentation	중·고등학생	c-rater™, c-rate-ML	지구과학	183명
18	Machine learning-enabled automated feedback: Supporting students' revision of scientific arguments based on data drawn from simulation	초·중학생	n-gram, SVM, 의사결정 나무	지구과학	1,034명
19	The impact of misspelled words on automated computer scoring: A case study of scientific explanations	대학생	품사 태깅, 명명된 개체 인식, NB, SVM	생명과학	-
20	Thematic analysis of 18 years of physics education research conference proceedings using natural language processing	물리교육 문헌	LDA	-	1,302편
21	Transforming biology assessment with machine learning: Automated scoring of written evolutionary explanations	대학생	SIDE	생명과학	565명
22	Utilizing a pretrained language model (BERT) to classify preservice physics teachers' written reflections	예비 물리 교사	BERT	물리학	-
23	Validation of automated scoring for a formative assessment that employs scientific argumentation	중학생	c-rate-ML	지구과학	1,180명
24	Validation of automated scoring of science assessments	-	c-rate-ML	물리학, 화학, 생명과학	-

-: 해당 논문에서 내용을 구체적으로 제시하지 않음

국내에서 이루어진 연구 동향은 Table 10과 같다.

연구 대상은 과학 교사, 대학생, 고등학생, 중학생 등에서 연구가 이루어졌고, 아직은 초등학생을 대상으로 연구가 이루어지지 않았다. 총 13편의 연구 중 7편이 문헌 분석 및 분류에 집중되었다는 점이 눈에 띈다. 우리나라의 연구에서 데이터를 수집하는 방법은 연구자가

직접 수업에서 데이터를 수집하거나 이미 작성된 문헌을 이용하는 방법이 있다. 이때 직접 데이터를 수집하는 방법은 비교적 많은 수를 구하기 어렵기에 문헌 분석 및 분류가 주를 이루는 것으로 보인다. 반면에, 해외에서는 교육 플랫폼이나 여러 교사가 협력하여 학생들의 텍스트 데이터를 대규모로 얻는 방법이 이루어지고 있기에 형성평가

Table 10. Domestic Research Trends in NLP

순번	제목	연구 대상	분석 프로그램	과학 내 과목	연구 대상 수
1	Doc2Vec을 이용한 중학교 과학 질문-답변 챗봇 개발 및 학생 질의 분석	중학생	형태소 분석, Doc2Vec	과학	질문 8,605개
2	과학 교사의 과학·AI융합교육 실천 동기와 인식	과학 교사	LDA	과학	15명
3	과학교육의 재미에 대한 재발견 -재미의 의미와 가치를 중심으로-	재미에 대한 국내 문헌	LDA	-	1,173편의 논문 초록
4	기계 학습을 활용한 논증 수준 자동 채점 및 논증 패턴 분석	고등학생	n-gram, SVM, 의사결정 나무, 랜덤 포레스트,	물리, 화학, 지구과학	353명
5	네트워크 텍스트 분석과 토픽모델링을 활용한 과학(영재)고 교육활동 내용분석	2021년 과학(영재)고 교육활동	텍스트 네트워크 분석, LDA	-	28개 과학(영재)고 교육 활동
6	머신 러닝을 활용한 과학 논문 구성 요소 코딩 자동화 가능성 탐색 연구	중학생	형태소 분석 및 태깅, TF-IDF, kNN, NB, 의사결정 나무, 인공신경망, SVM	생명과학	중학교 1학년 37명 중학교 2학년 30명
7	머신러닝을 활용한 영재 학생들의 졸업논문 주제 경향 분석	영재학교 졸업논문	BERT, kNN	-	292편
8	온라인 과학 기사 텍스트 마이닝을 통해 분석한 에너지 용어 사용의 맥락	과학 뉴스	언어 네트워크 분석, 주제 분석	물리	기사 2,487
9	잠재 디리클레 할당(LDA) 기반의 토픽모델링 분석을 통한 ‘초등과학교육’ 학술지 연구논문의 주제 및 변화	초등과학교육 문헌	LDA	-	1,065편
10	전문가의 형태소 분류를 활용한 과학 논증 자동 채점	고등학생	토큰화, 형태소 분석, 품사 태깅, 문서단어행렬, SVM, 의사결정 나무, 랜덤 포레스트, 인공신경망	과학관련 사회쟁점	1학년 250명 2학년 103명
11	텍스트마이닝을 활용한 대학생들의 외계행성 개념 변화 연구	대학생	n-gram	지구과학	30명
12	토픽 모델링을 활용한 과학영재교육 연구동향 분석	과학영재교육 문헌	LDA	-	292편
13	한국과학교육학회지는 44년간 어떤 주제로 어떻게 변화했는가? -잠재 디리클레 할당(LDA)을 활용한 토픽모델링 분석	과학교육 문헌	LDA	-	2,115편의 논문 초록

및 자동 채점 분야의 연구가 활발히 진행되었다. 다만, 해외에서 진행된 연구에서도 대규모 데이터를 수집하는 방법의 어려움을 지적하고 있다. 따라서 국내에서 과학교육에 NLP 관련 연구가 활발하게 진행되기 위해서는 국내 연구자들이 NLP 기법의 필요성을 인지하고, 연구에 필요한 데이터를 쉽게 수집할 수 있는 환경 조성이 필요하다.

V. 결론 및 제언

이 연구에서는 과학교육 문헌에서 NLP 기법의 주요 응용 분야, NLP 기법을 활용할 때 교사의 역할을 분석하였고, 국내외에서 이루어진 NLP 기법을 활용한 과학교육 문헌을 비교하였다. 연구 결과를 종합해 보면 다음과 같다.

첫째, 과학교육 분야에서 형성평가, 자동 채점, 문헌 검토 및 분류, 그리고 패턴 추출 등의 주요 분야에서 NLP 기법의 활용이 증가하고 있고, 그 효과가 두드러지게 나타나고 있다. NLP 기법을 통해 학생들의 학습 과정과 이해도를 실시간으로 파악하여 학생들에게 더욱 정확하고 효과적인 피드백을 제공할 수 있다. 또한 NLP 기법을 활용하면 대량의 문헌 데이터를 효과적으로 분석할 수 있으며, 연구의 주제나 범주를 정확하게 파악할 수 있다. 그뿐만 아니라 NLP 기법을 활용해 학생들의 응답이나 설명들에서 나타나는 공통점을 찾거나 패턴을 추출하여 유형을 분류할 수 있다. 하지만 지금의 NLP 기법은 학생들의 복잡한 사고와 깊은 이해를 완전히 파악하는데 한계가 있기 때문에 더욱 정교한 NLP 기법이 개발되어야 하고, 대량의 고품질 훈련 데이

터가 필요하지만 이를 확보하기 어려운 문제점도 존재한다.

둘째, 과학교육에서 NLP 기법을 도입하여 교사의 역할이 단순한 지식 전달자에서 학생들의 학습을 지원하고 촉진하는 지도자로 확장되었다. NLP 기법을 활용해 사들은 학생들의 학습 과정과 결과를 더욱 깊게 이해하고, 그에 따라 교육 방법과 교육 내용을 개선할 수 있게 되었다. 또한 NLP 기법을 활용한 교육은 학생들에게도 다양한 이점을 제공한다. 학생들은 즉각적인 피드백을 받을 수 있어 학습의 효과성을 높일 수 있으며, 자기주도 학습을 통해 학습의 주체성을 강화할 수 있다. 그러나 과학교육에서 NLP를 성공적으로 도입하기 위해 교사들은 NLP와 기계학습을 효과적으로 분석하고 해석할 수 있는 능력과 새로운 기술 개발을 따라가기 위해 지속적인 전문성을 개발해야 한다. 그리고 NLP 도구의 개선을 위해서 NLP 개발자에게 피드백을 제공하고 과학 교육과정 및 교육 목표와의 통합을 위해 노력해야 한다.

셋째, 2011년부터 2023년 9월까지의 연구 동향을 살펴보면, 국내에서 NLP 기법의 활용 연구는 주로 문헌 검토 및 분류에 집중되고 있다. 해외에서는 교육 플랫폼 및 공동 연구를 통해 대규모 데이터를 수집하는 데 비해 국내에서는 대규모 데이터를 수집하는 방법을 겪기 때문에 기존의 문헌을 활용하는 연구가 주를 이루는 것으로 보인다. 따라서 국내에서 NLP가 다양한 분야에서 활발하게 연구되기 위해서는 데이터를 쉽게 수집할 수 있는 환경 조성이 필요하다.

본 연구의 결론을 토대로 과학교육에서 NLP 기법의 활용과 관련하여 다음과 같은 제언을 하고자 한다.

첫째, 교사를 대상으로 NLP 기법의 기본 원리를 익히고, NLP 기법을 활용해 교육 프로그램을 개발하여 활용하는 방법에 대한 연수를 진행할 필요가 있다. 과학교육 분야에서 NLP 기법의 활용은 교육의 효과를 향상할 수 있는 중요한 도구로 간주하고 있으며, 교사와 학생 모두에게 긍정적인 변화를 불러온다고 보고되고 있다. 따라서 교사들이 NLP 기법에 대한 학습을 통해 과학교육 상황에서 NLP 기법을 효과적으로 사용하고, 수업 및 평가의 질을 향상할 수 있을 것이다.

둘째, 과학교육에서 NLP 기법의 활용과 관련한 연구가 지속해서 이루어져야 한다. 국내의 NLP 관련 연구는 문헌 검토 및 분류에 집중되어 있다. 이를 개선하기 위해서는 국내의 언어적 및 문화적 배경을 고려한 NLP 기법의 개발이 필요하다. 특히 한국어의 구문과 의미 체계를 이해하고 이를 반영한 NLP 모델을 개발하여 학생들의 학습 데이터를 분석하는 것이 중요하다. 이러한 접근 방식은 과학교육의 질을 높이고, 학습 효과를 개선하는 데 기여할 것이다.

셋째, 과학교육에 NLP 기법을 효과적으로 적용하기 위해서는 체계적인 데이터 관리가 필요하다. 해외 연구 사례에서 볼 수 있듯이 다양하고 방대한 데이터 활용은 NLP 연구의 품질을 크게 향상할 수 있다. 따라서 국내에서도 텍스트 데이터에 대한 접근성과 활용도를 높이는 데 중점을 두어야 한다. 이를 위해 적절한 데이터 관리 및 분석 인프라 구축이 필요하며, 이를 통해 다양한 분야의 텍스트 데이터를 깊이 있게 분석하고, 그 결과를 과학교육과 같은 실제적인 응용 분야에 활용할 수 있을 것이다.

국문요약

본 연구는 2011년부터 2023년 9월까지 과학교육 분야에서 자연어 처리(NLP) 기법을 적용한 37건의 국내 및 해외 문헌을 분석하여 과학교육에서의 NLP 관련 연구 동향을 파악하고자 하였다. 특히 과학교육에서 NLP 기법의 주요 응용 분야, NLP 기법을 활용할 때 교사의 역할, 국내외 해외의 비교 측면에서 그 내용을 체계적으로 분석하였다. 분석 결과는 다음과 같다. 첫째, NLP 기법이 과학교육에서 형성평가, 자동 채점, 문헌 검토 및 분류, 패턴 추출에 중요하게 활용되고 있음을 확인하였다. 형성평가에서 NLP를 활용하면 학생들의 학습 과정과 이해도를 실시간으로 분석할 수 있다. 이는 교사의 수업에 대한 부담을 줄이고, 학생들에게 정확하고 효과적인 피드백을 제공할 수 있다. 자동 채점에서는 학생들의 응답을 빠르고 정확하게 평가하는 데 기여한다. 문헌 검토 및 분류에서는 과학교육 관련 연구나 학생들의 보고서를 분석하여 주제와 트렌드를 효과적으로 분석하고, 미래 연구 방향을 설정하는 데 도움을 준다. NLP 기법을 패턴 추출에 활용하면 학생들의 생각과 반응에 나타난 공통점이나 패턴을 찾아 효과적으로 분석할 수 있다. 둘째, 과학교육에서 NLP 기법의 도입은 교사의 역할을 지식 전달자에서 학생들의 학습을 지원하고 촉진하는 지도자로 확장했고, 교사들에게 지속적인 전문성 개발을 요구한다. 셋째, 국내에서는 문헌 검토 및 분류에 집중되어 있어 국내 NLP 연구의 다양성을 위해 텍스트 데이터 수집이 용이한 환경 조성이 필요하다. 이러한 분석 결과를 바탕으로 과학교육에서 NLP 기법의 활용하는 방법에 대해 논의하였다.

주제어 : 자연어 처리, 과학교육, 교사의 역할

References

- Beggrow, E. P., Ha, M., Nehm, R. H., Pearl, D., & Boone, W. J. (2014). Assessing scientific practices using machine-learning methods: How closely do they match clinical interview performance?. *Journal of Science Education and Technology*, 23, 160-182.
- Brown, P. F., Della Pietra, V. J., Desouza, P. V., Lai, J. C., & Mercer, R. L. (1992). Class-based n-gram models of natural language. *Computational Linguistics*, 18(4), 467-480.
- Chang, J., & Na, J. (2022). An examination of the topics and changes in the research papers published in the *Journal of Korean Elementary Science Education* using latent dirichlet allocation for the topic modeling analysis. *Journal of Korean Elementary Science Education*, 41(2), 356-372.
- Chang, J., & Na, J. (2022). How the *Journal of the Korean Association for Science Education*(JKASE) changed for the past 44 years?: Topic modeling analysis using latent dirichlet allocation. *Journal of the Korean Association for Science Education*, 42(2), 185-200.
- Choi, J., Song, H., & Nam, K. (2010). Formulaic expressions in Korean. *Discourse and Cognition*, 17(2), 163-190.
- Chowdhary, K., & Chowdhary, K. R. (2020). Natural language processing. *Foundamentals of artificial intelligence*, 603-649.
- Deng, L., & Yu, D. (2014). Deep learning: methods and applications. *Foundations and trends® in signal processing*, 7(3-4), 197-387.
- Donnelly, D. F., Vitale, J. M., & Linn, M. C. (2015). Automated guidance for thermodynamics essays: Critiquing versus revisiting. *Journal of Science Education and Technology*, 24, 861-874.
- Gombert, S., Di Mitri, D., Karademir, O., Kubsch, M., Kolbe, H., Tautz, S., Grimm, A., Bohm, I., Neumann, K., & Drachler, H. (2023). Coding energy knowledge in constructed responses with explainable NLP models. *Journal of Computer Assisted Learning*, 39(3), 767-786.
- Ha, M., & Nehm, R. H. (2016). The impact of misspelled words on automated computer scoring: A case study of scientific explanations. *Journal of Science Education and Technology*, 25, 358-374.
- Ha, M., Nehm, R. H., Urban-Lurain, M., & Merrill, J. E. (2011). Applying computerized-scoring models of written biological explanations across courses and colleges: Prospects and limitations. *CBE-Life Sciences Education*, 10(4), 379-393.
- Han, S., Kim, Y., & Kim, H. (2020). A study on the conceptual changes of extra-solar planet in university students using text-mining techniques. *Journal of Korean Society of Earth Science Education*, 13(3), 305-316.
- Isotani, S., Millán, E., Ogan, A., Hastings, P., McLaren, B., & Luckin, R. (2019). *Artificial intelligence in education*. Chicago: Springer International Publishing.
- Jho, H. (2023). Understanding of generative artificial intelligence based on textual data and discussion for its application in science education. *Journal of the Korean Association for Science Education*, 43(3), 307-319.
- Jho, H., & Lee, B. (2022). Clustering science gifted students' graduation theses based on machine learning. *Journal of Science Education for the Gifted*, 14(1), 13-22.
- Jockers, M. L., & Thalken, R. (2020). *Text analysis with R*. New York: Springer International Publishing.
- Kang, M., Chaudhuri, S., Joshi, M., & Rose, C. (2008). Side: The summarization integrated development environment. In *Proceedings of the ACL-08: HLT Demo Session*, 24-27.
- Kim, H., & Jhun, Y. (2021). Analysis of trends in science gifted education using topic modeling. *Journal of Korean Elementary Science Education*, 40(3), 283-294.
- Kim, J., & Oh, Y. (2023). Content analysis of education activities of science high schools using network text analysis and topic modeling. *Journal of Gifted/Talented Education*, 33(3), 375-402.
- Lantz, B. (2019). *Machine learning with R: expert techniques for predictive modeling*. UK: Packt publishing ltd.
- Lee, G., Ha H., Hong, H., & Kim, H. (2018). Exploratory research on automating the analysis of scientific argumentation using machine learning. *Journal of the Korean Association for Science Education*, 38(2), 219-234.
- Lee, H. S., Gweon, G. H., Lord, T., Paessel, N., Pallant, A., & Pryputniewicz, S. (2021). Machine learning-enabled automated feedback: Supporting students' revision of scientific arguments based on data drawn from simulation. *Journal of Science Education and Technology*, 30, 168-192.
- Lee, H. S., Pallant, A., Pryputniewicz, S., Lord, T., Mulholland, M., & Liu, O. L. (2019). Automated text scoring and real-time adjustable feedback: Supporting revision of scientific arguments involving uncertainty. *Science Education*, 103(3), 590-622.
- Lee, M., & Ryu, S. (2020). Automated scoring of scientific argumentation

- using expert morpheme classification approaches. *Journal of the Korean Association for Science Education*, 40(3), 321-336.
- Lee, M., & Ryu, S. (2021). Automated scoring of argumentation levels and analysis of argumentation patterns using machine learning. *Journal of the Korean Association for Science Education*, 41(3), 203-220.
- Lim, H. (2019). *Natural language processing bible*. Seoul: Human Science.
- Litman, D. (2016). Natural language processing for enhancing teaching and learning. In *Proceedings of the AAAI conference on artificial intelligence*, 30(1), 4170-4176.
- Liu, O. L., Brew, C., Blackmore, J., Gerard, L., Madhok, J., & Linn, M. C. (2014). Automated scoring of constructed-response science items: Prospects and obstacles. *Educational Measurement: Issues and Practice*, 33(2), 19-28.
- Liu, O. L., Rios, J. A., Heilman, M., Gerard, L., & Linn, M. C. (2016). Validation of automated scoring of science assessments. *Journal of Research in Science Teaching*, 53(2), 215-233.
- Mao, L., Liu, O. L., Roohr, K., Belur, V., Mulholland, M., Lee, H. S., & Pallant, A. (2018). Validation of automated scoring for a formative assessment that employs scientific argumentation. *Educational Assessment*, 23(2), 121-138.
- Michaud, L. N., & McCoy, K. F. (2006). Capturing the evolution of grammatical knowledge in a CALL system for deaf learners of English. *International Journal of Artificial Intelligence in Education*, 16(1), 65-97.
- Min, G., & Yoo, J. (2022). Development of a middle school science Q&A chatbot using Doc2Vec and analysis of student's queries. *The SNU Journal of Education Research*, 31(3), 115-145.
- Nakamura, C. M., Murphy, S. K., Christel, M. G., Stevens, S. M., & Zollman, D. A. (2016). Automated analysis of short responses in an interactive synthetic tutoring system for introductory physics. *Physical Review Physics Education Research*, 12(1), 010122.
- Nehm, R. H., Ha, M., & Mayfield, E. (2012). Transforming biology assessment with machine learning: automated scoring of written evolutionary explanations. *Journal of Science Education and Technology*, 21, 183-196.
- Odden, T. O. B., Marin, A., & Caballero, M. D. (2020). Thematic analysis of 18 years of physics education research conference proceedings using natural language processing. *Physical Review Physics Education Research*, 16(1), 010142.
- Odden, T. O. B., Marin, A., & Rudolph, J. L. (2021). How has Science education changed over the last 100 years? An analysis using natural language processing. *Science Education*, 105(4), 653-680.
- Oh, C., & Kang, N. (2021). Analyzing different contexts for energy terms through text mining of online science news articles. *Journal of Science Education*, 45(3), 292-303.
- Park, C., Kim, Y., Kim, J., Song, J., & Choi, H. (2015). *R data mining*. Seoul: Kyowoo.
- Reese, R. M., & Bhatia, A. (2018). *Natural language processing with Java: Techniques for building machine learning and neural network models for NLP*. Birmingham: Packt Publishing Ltd.
- Rosenberg, J. M., & Krist, C. (2021). Combining machine learning and qualitative methods to elaborate students' ideas about the generality of their model-based explanations. *Journal of Science Education and Technology*, 30, 255-267.
- Sarkar, D. (2019). *Text analytics with Python: a practitioner's guide to natural language processing*. Bangalore: Apress.
- Shaik, T., Tao, X., Li, Y., Dann, C., McDonald, J., Redmond, P., & Galligan, L. (2022). A review of the trends and challenges in adopting natural language processing methods for education feedback analysis. *IEEE Access*, 10, 56720-56739.
- Sherin, B. (2013). A computational study of commonsense science: An exploration in the automated analysis of clinical interview data. *Journal of the Learning Sciences*, 22(4), 600-638.
- Shin, E. (2022). Science teachers' motivation and perception of science · AI convergence education. *The Korean Society for School Science*, 16(3), 398-412.
- Shin, S., Ha, M., & Lee, J. (2018). Rediscovering the interest of science education: Focus on the meaning and value of interest. *Journal of the Korean Association for Science Education*, 38(5), 705-720.
- Sung, S. H., Li, C., Chen, G., Huang, X., Xie, C., Massicotte, J., & Shen, J. (2021). How does augmented observation facilitate multimodal representational thinking? Applying deep learning to decode complex student construct. *Journal of Science Education and Technology*, 30, 210-226.
- Tili, A., Shehata, B., Adarkwah, M. A., Bozkurt, A., Hickey, D. T., Huang, R., & Agyemang, B. (2023). What if the devil is my guardian angel: ChatGPT as a case study of using chatbots in education. *Smart Learning Environments*, 10(1), 15.
- Wilson, J., Pollard, B., Aiken, J. M., Caballero, M. D., & Lewandowski, H. J. (2022). Classification of open-ended responses to a research-based assessment using natural language processing. *Physical Review Physics Education Research*, 18(1), 010141.
- Wulff, P., Buschhüter, D., Westphal, A., Mientus, L., Nowak, A., & Borowski, A. (2022). Bridging the gap between qualitative and quantitative assessment in science education research with machine learning-A case for pretrained language models-based clustering. *Journal of Science Education and Technology*, 31(4), 490-513.
- Wulff, P., Buschhüter, D., Westphal, A., Nowak, A., Becker, L., Robalino, H., Steda, M., & Borowski, A. (2021). Computer-based classification of preservice physics teachers' written reflections. *Journal of Science Education and Technology*, 30, 1-15.
- Wulff, P., Mientus, L., Nowak, A., & Borowski, A. (2023). Utilizing a pretrained language model (BERT) to classify preservice physics teachers' written reflections. *International Journal of Artificial Intelligence in Education*, 33(3), 439-466.
- Wulff, P., Westphal, A., Mientus, L., Nowak, A., & Borowski, A. (2023). Enhancing writing analytics in science education research with machine learning and natural language processing-Formative assessment of science and non-science preservice teachers' written reflections. In *Frontiers in Education*. 7, 1061461.
- Yoo, J. E. (2019). Machine learning for large-scale/panel data and learning analytics data analysis. *Journal of Educational Technology*, 35(2), 313-338.
- Zhai, X., He, P., & Krajcik, J. (2022). Applying machine learning to automatically assess scientific models. *Journal of Research in Science Teaching*, 59(10), 1765-1794.
- Zhu, M., Lee, H. S., Wang, T., Liu, O. L., Belur, V., & Pallant, A. (2017). Investigating the impact of automated feedback on students' scientific argumentation. *International Journal of Science Education*, 39(12), 1648-1668.

저자정보

전철홍(한국교육대학교 학생)
유선아(한국교육대학교 교수)