

# 데이터 리터러시를 위한 머신러닝 기반 AI 융합 수업 모형 개발

강상우, 이유진\*, 임효정, 최원근  
서울대학교 AI융합교육학과 석사과정 학생

## Development of AI Convergence Education Model Based on Machine Learning for Data Literacy

Sang-Woo Kang, Yoo-Jin Lee\*, Hyo-Jeong Lim, Won-Keun Choi  
Student, Department of AI Convergence Education, Seoul National University

**요약** 본 연구는 고등학교 학생들의 데이터 리터러시를 함양할 수 있는 머신러닝 기반 AI 융합 수업 모형과 수업 설계 원리를 개발하고, 그에 따른 상세 지침을 개발하는 것을 목적으로 하였다. 이를 위해 선행 문헌 연구를 통해 머신러닝을 기반으로 한 수업 모형과 설계 원리 및 상세 지침을 개발하고, 서울 소재 상업계열 특성화고등학교 학생 15명에게 적용하여 실행하였다. 연구 결과 학생들의 데이터 리터러시가 통계적으로 유의미( $p < .001$ )하게 향상되었으므로 본 연구의 수업 모형이 학습자의 데이터 리터러시 향상에 긍정적인 영향을 주었음을 확인할 수 있었고, 앞으로 관련 연구로 이어지길 기대한다.

**주제어** : 머신러닝, 인공지능, 융합교육, 데이터 리터러시, 결정트리

**Abstract** The purpose of this study is to develop a machine learning-based AI convergence class model and class design principles that can foster data literacy in high school students, and to develop detailed guidelines accordingly. We developed a machine learning-based teaching model, design principles, and detailed guidelines through research on prior literature, and applied them to 15 students at a specialized high school in Seoul. As a result of the study, students' data literacy improved statistically significantly ( $p < .001$ ), so we confirmed that the model of this study has a positive effect on improving learners' data literacy, and it is expected that it will lead to related research in the future.

**Key Words** : Machine Learning, Artificial intelligence, Convergence Education, Data Literacy, Decision Tree

### 1. 서론

해외 각 나라에서는 교육분야에서 인공지능 인재를 양성하기 위한 다양한 정책을 제시하고 있다. 미국의 경우 인공지능발전협회(AAAI : Association for the Advancement of Artificial Intelligence)와 컴퓨터과

학 교사협회(CSTA : Computer Science Teachers Association)에서 AI4K12라는 교육과정을 만들어 5 가지 핵심 아이디어를 중심으로 가이드라인을 제공하였다[1]. 초기에는 컴퓨터과학을 중심으로 인공지능 교육을 실시했지만 인공지능을 각 교과에 도입해야한다

\*Corresponding Author : Yoo-Jin Lee(yoo512@snu.ac.kr)

는 연구 결과가 도출되면서 언어, 수학, 과학, 사회 등 다양한 교육과정에 인공지능에 대한 문해력을 추가하려는 노력이 지속되고 있다[2].

우리나라 또한 2020년 '인공지능시대 교육정책방향과 핵심과제'를 발표하면서 인공지능 교육에 대한 비전을 제시하였다[3]. 특히 '2022개정 교육과정'에서는 인공지능과 융합한 다양한 교과를 개발하여 학생들이 인공지능을 다양한 교과와 융합하여 학습할 수 있도록 기회를 제공하고 있다[4]. 또한 실제 교육현장에 인공지능과 교과가 융합된 수업이 이루어질 수 있도록 교육대학원에 '인공지능 융합교육대학원'을 개설하여 인공지능 융합교육 역량을 지닌 교원들을 양성하고 인공지능 융합교육 관련 연구를 수행하고 있다[3]. 이외에도 인공지능 교육을 발전시키기 위한 교육계의 노력은 다양하게 이루어지고 있다.

인공지능을 활용한 교육 프로그램은 학생들의 인공지능에 대한 인식, 인공지능에 대한 기본 소양과 역량 등을 향상시킬 수 있다[5]. 특히 머신러닝의 원리를 학생들의 눈높이에 맞게 구현한 다양한 플랫폼들이 개발됨에 따라 머신러닝을 비교적 흥미롭게 학습할 수 있고, 이를 통해 인공지능과 관련된 다양한 역량을 함양할 수 있었다. 이외에도 김선민, 박기범(2023)은 머신러닝의 알고리즘을 언플러그드 수업으로 구성하였고 그 결과 학생들의 정보활용능력이 향상됨을 확인하였다[6]. 이처럼 머신러닝은 인공지능을 이해하는 가장 기본적인 원리로서 학생들에게 인공지능의 유용성을 인지시키는데 좋은 도구가 될 수 있다.

기본적으로 인공지능 기술은 빅데이터를 기반으로 알고리즘을 개발한 것이다[7]. 따라서 인공지능을 이해하기 위해서는 인공지능에 사용되는 데이터를 다루고 이해하는 역량이 필요하다. 또한 Wolff et al.(2016)에 따르면 데이터 리터러시의 향상은 일반적으로 머신러닝을 활용하는 능력을 길러준다[8]. 이러한 데이터와 관련된 역량을 포괄하는 개념으로 주목받는 것이 데이터 리터러시이다. 다만, 데이터 리터러시는 단순히 데이터에 대한 학습 능력을 넘어서 데이터를 활용하여 자신의 주장을 뒷받침하고 합리적인 의사결정을 하는 능력이기 때문에 단순한 데이터 분석 기술 이상의 수업이 필요하다[9]. 데이터 리터러시 역량은 이미 다양한 분야에서 연구되고 있으며 이를 함양하기 위한 수업모형과 교수전략이 개발되고 있다. 송유경(2021)은 데이터 리터

러시를 함양하기 위한 토론 수업 모형을 개발하여 학생들의 데이터 리터러시 함양에 유의미한 향상을 발견하였고, 조연수(2022)는 인공지능을 융합한 과학 수업을 통해 학생들의 데이터 리터러시 역량이 유의미하게 증가함을 발견하였다[10].

머신러닝을 기반으로 한 수업은 다양한 데이터를 수집, 분석하고 이들을 머신러닝을 통해 분류하고 처리하면서 학생들의 데이터 리터러시 능력을 향상시킬 수 있다. 고등학교 수준에서는 수준 높은 머신러닝의 구현은 어렵지만 엔트리 등의 프로그램을 활용하면 충분히 머신러닝을 구현할 수 있으며, 이러한 머신러닝 기법을 바탕으로 실생활 문제를 다루면 보다 능동적인 문제해결활동이 이루어질 수 있어 수업의 효과가 증대될 수 있다[11]. 본 연구에서는 머신러닝 기법을 바탕으로 학생들의 데이터 리터러시를 향상시키는 AI 융합 수업 모형을 개발하고, 타당성을 검증하고자 한다.

## 2. 관련 연구

### 2.1 데이터와 인공지능

데이터에 대한 이해가 없으면 인공지능을 제대로 이해할 수 없으며, 데이터를 활용하려면 다양한 수준의 인공지능 지식이 필요하다[12]. Fig. 1은 인공지능과 관련된 개념들 사이의 관계를 도식화한 것으로 데이터와 인공지능이 밀접하게 관련되어 있음이 나타나 있다.

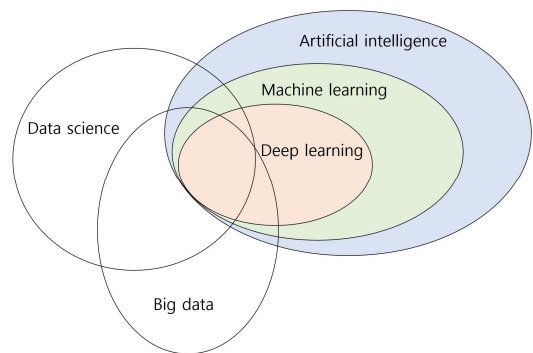


Fig. 1. Inclusion relationships between AI-related concepts

각국의 인공지능 관련 교육 지침에도 데이터와 관련된 내용이 포함되어 있다. 우리나라 정보 교과 2022 개정 교육과정에서는 인공지능 소양을 인공지능 윤리의식과 데이터에 대한 이해를 기반으로 인공지능을 통

해 문제를 해결할 수 있는 능력이라 정의하고 인공지능 소양의 하위 역량 중 하나로 데이터 문해력을 제시하고 있다[4]. 미국 AAI와 CSTA가 공동 설립한 AI4K12 이니셔티브는 5 Big Idea in AI에 데이터를 통한 학습을 세 번째 항목으로 포함하고 그 세부 내용으로 머신러닝이 데이터의 패턴을 찾는 일종의 통계적 추론이며 인공지능의 발전에 많은 양의 데이터가 필요함을 Fig. 2와 같이 제시하였다[15]. 캐나다의 STEM(Science, Technology, Engineering, Mathematics) 교육 지원 단체 Actua는 AI4K12의 다섯 가지 항목에 데이터 테마를 추가하여 AI 교육 프레임워크를 6가지 테마로 설정하고 그 중 학습 테마에서 머신러닝이 데이터로 이루어진다는 것을 Big Idea로 제시하였다[16].

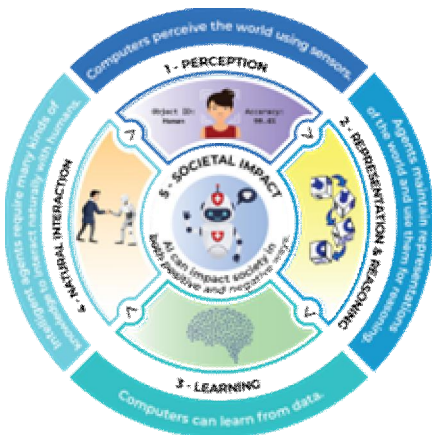


Fig. 2. The Five Big Idea of AI (AI4K12, 2021)

Long, Magerko(2020)는 인공지능 역량을 17가지로 정리하면서 그 중 세 가지를 데이터 리터러시, 데이터로부터의 학습, 데이터를 비판적으로 해석하기 등 데이터와 관련된 항목으로 선정하였으며 인공지능 교육 디자인 시 고려 요소에도 데이터 맥락화를 추가하였다[17]. Olari, Romeike(2021)는 인공지능 역량을 기르기 위한 대부분의 커리큘럼이 데이터 다양성 유지, 효율적 분류, 목표 달성을 위한 데이터셋 식별 등 데이터에 관련한 내용을 포함하고 있으면서도 데이터 리터러시에 대해 종합적으로 다루고 있지는 않음을 지적하였다. 이에 더해 데이터 수명 주기를 활용하여 인공지능과 데이터 리터러시를 종합적으로 고려한 교육과정을 수립해야 한다고 주장하였다[12].

## 2.2 데이터 리터러시

데이터 리터러시는 데이터를 통계적으로 다루는 능력과 데이터를 선별하고 이를 활용하여 다양한 문제를 해결하는 능력을 포함한다. 데이터 리터러시는 연구의 맥락에 따른 구성요소들을 포함하고 있어 연구자마다 조금씩 다르게 사용하고 있다[18]. Table 1은 각 학자들의 데이터 리터러시에 대한 정의이다.

Table 1. Definition of data literacy

Definition of data literacy
Ability to analyze data and use it for purpose. The ability to discover, manipulate, interpret, visualize, and communicate data, including numbers as well as text and images [19].
Ethically considering the use of data, the ability to raise questions and provide answers to real-life problems through the collection of both large and small datasets [18].
The competence involving digital technical skills such as understanding, collecting, preprocessing, analyzing and interpreting data, visualizing information, and preserving data, combined with the humanities-based thinking and logic required for their appropriate utilization [20].
The series of processes involving extracting information from data, generating questions about data, utilizing data for real-world problem-solving, organizing data using appropriate tools, and drawing conclusions, all based on data [21].
The competency demanded in the contemporary era, where the utilization of data is increasing, involves the ability to interpret data for its purposeful application. This includes skills in data collection, management, processing and analysis, visualization, as well as planning abilities [22].

2022 교육과정에 신설된 ‘데이터 과학’ 교과는 컴퓨팅 사고력을 기반으로 디지털 사회에서 데이터의 역할 및 잠재적 가치와 데이터 과학에 기반한 문제 해결 과정의 중요성을 인식하며, 다양한 분야의 문제를 해결하고 합리적 의사 결정을 위한 통찰의 역량을 키우는 것을 목표로 한다[4]. 이로 미루어 보았을 때, 우리나라 교육과정에서는 데이터와 데이터에 기반한 문제 해결의 가치를 인식하는 것을 데이터 리터러시의 일부로 다루고 있음을 알 수 있다.

본 연구에서는 다양한 구성 요소 중 머신러닝과 밀접하게 연관된 요소를 선택하여 데이터 리터러시를 데이터를 이해, 수집 및 분석하고 데이터를 기반으로 의사소통하는 능력이라 정의하였다.

### 2.3 머신러닝

머신러닝(Machine Learning, 기계학습)은 컴퓨터가 사전에 프로그래밍되어 있지 않고, 데이터로부터 학습하여 새로운 데이터에 대해 적절한 작업을 수행하는 일련의 알고리즘이나 처리 과정을 의미한다. 머신러닝은 인공지능의 하위 범주에 속하며, 컴퓨터 과학, 통계학, 데이터 마이닝 등의 분야와 밀접한 관련이 있다[23].

머신러닝은 학습 방식에 따라 지도학습, 비지도학습, 강화학습이 있다. 데이터를 사용하여 훈련시킨 모델을 이용하여 예측하는 지도학습은 학습 데이터에 레이블이 있어 머신러닝 모델을 학습시키는 방식이다. 지도학습은 문제 형식에 따라 분류(Classification)와 예측(Regression 또는 Prediction) 모델로 나눌 수 있다. 비지도학습은 학습 데이터에 레이블이 없으며, 주로 입력하는 변수들 사이의 관계를 이용하여 분류하는 군집(Clustering) 모델이 있다. 마지막으로 강화학습은 알고리즘이 수행한 결과에 따라 수행 방식을 진화시켜 나가는 머신러닝 모델이다. 주어진 환경에서 인공지능이 상태를 인식하고 행동한 결과에 따른 보상으로 최적화하며 스스로 시행착오를 겪으며 학습하고 진화한다는 점에서 지도학습이나 비지도학습과 큰 차이를 보인다[24].

머신러닝은 정보 시스템이 일상의 거의 모든 분야에 파고든 현 사회에서 급속히 발전하고 있는 중요한 분야로서 학생들에게도 그들의 컴퓨팅 실력에 맞추어 머신러닝 교육을 실시해야 한다[25]. 김동용 외(2023)는 초등학교의 문제해결력 향상을 위한 머신러닝 교육 프로그램을 개발하였다. 초등학교가 원리와 개념을 이해하고 흥미를 가지고 학습할 수 있도록 AI for Oceans, 퀵드로우(Quick Draw), 오토드로우(Auto Draw), 구글두들바흐(Google Doodle Bach)를 도구로 사용하였으며 프로그램 실시 후 문제해결력이 통계적으로 유의미하게 향상하였다[26]. 이는 초등학교에게도 적절한 도구를 사용하면 머신러닝을 활용한 학습이 가능함을 보여준다.

문우중 외(2021)에 따르면 머신러닝에 필요한 데이터를 수집하고 적절하게 분류하는 데이터 라벨링은 머신러닝 학습의 중요한 사전작업이며 머신러닝의 성능에 영향을 미친다[27]. 머신러닝 모델을 학습시키기 위해서는 데이터가 필요하고 모델의 성능을 향상시키기 위해서는 데이터 이해, 수집, 분석 등이 필수적이며 이는 데이터 리터러시의 요소와 일치한다. 이를 바탕으로

볼 때, 머신러닝 교육으로 학생들의 데이터 리터러시를 키우고 데이터 수집 및 분석 능력을 향상시킬 수 있을 것으로 기대할 수 있다.

### 2.4 인공지능 융합 교육

인공지능 융합 교육은 인공지능 기술, 인공지능에 대한 원리와 핵심 개념 이해를 기반으로 다양한 학문(교과)간 및 산업분야를 융합하여 창의적인 해결책을 제시하는 교육이다[28]. 인공지능 융합 교육은 인공지능 자체에 대한 이해를 목적으로 하는 '인공지능 이해 교육', 인공지능을 도구로 활용하는 교육인 '인공지능 활용 교육'과 구분된다[29].

일반교과에서 융합 교육을 실시하는 것이 쉽지 않음을 STEAM 교육의 현황을 통해 알 수 있다. 권혁수 외(2021)에 따르면 2011년부터 교육부가 융합인재교육(STEAM)을 제안하고 각각도로 지원하고 있음에도 2020년 실시된 STEAM 교육 운영 현황 실태조사 결과에 따르면 응답자 중 STEAM 교육을 실행하고 있는 교사는 초등학교 59.3%, 중학교 42.1%, 고등학교 35.2%에 불과했다[30]. 융합교육을 미실시하는 이유로는 시험 준비로 인한 교육과정 재구성의 어려움, 정규 교과에서 다른 교과와의 협업 진행의 어려움, 정해진 수업 시수 조정의 어려움 등이 있었다. 더구나 인공지능 융합교육이 가능하기 위해서는 학생들이 인공지능을 활용할 수 있어야 한다. 즉 학생들에게 인공지능 기술 사용법을 가르쳐야 한다는 점 또한 인공지능 융합 교육을 시도하는 것에 대한 장애물이 될 수 있다.

이와 같은 어려움에도 불구하고 인공지능 조작 경험, 인공지능 교육경험이 많은 학생들이 인공지능에 대한 긍정적 태도가 높게 나타난 연구 결과는 일반 교과의 다양한 상황에서 인공지능 융합교육을 시도해야 함을 보여준다[31]. 일반 교과와 관련된 문제들을 소프트웨어적 관점에서 접근하고 해결하기 위해서는 소프트웨어가 독립된 교과목으로서만 존재하는 것보다 전 교과목들과 유기적 관련성을 갖도록 편성, 운영될 필요가 있다[32].

최근 고등학생들을 대상으로 일반 교과에서 인공지능을 융합하여 데이터 리터러시를 기르기 위한 연구가 수행되고 있다. 송유경(2021)은 데이터 리터러시 향상을 위한 데이터 기반 토론 수업 모형을 개발하였고[9], 노동규(2023)는 데이터 리터러시 향상을 위한 공개된

과학 데이터셋 기반의 회귀 모델을 만드는 수업 모형을 개발하였으며 채승철, 조정효, 이준행(2021)은 구글 코랩과 PyTorch를 이용한 과학 데이터 기반 교육 자료를 개발하였다[33,34].

그러나 기존 연구들은 수업에 많은 시수를 사용하거나 텍스트 코딩을 배우지 않았다면 사용이 쉽지 않은 도구를 이용하여 현장 적용에 어려움이 있다. 고등학교에서 수업시수의 부담을 줄이고 일반 교과에 적용할 수 있도록 도구에 대한 학습 부담이 적고 수업 단계를 단순화하여 교육활동에 비교적 적은 차시를 사용하는 수업 모형 개발이 필요하다.

일반 교과에서의 교과교육활동에서는 종종 데이터가 생성되며 인공지능 시스템은 데이터를 필요로 하므로 이러한 교과교육활동의 데이터를 인공지능 융합 지점으로 활용할 수 있다[2]. 따라서 본 연구에서는 일반 교과 학습 주제와 관련된 데이터를 이용한 머신러닝을 통해 데이터 리터러시를 기를 수 있는 수업 모형을 제시하고자 한다.

### 3. 연구 방법

#### 3.1 설계, 개발 연구

본 연구는 데이터 리터러시를 위한 머신러닝 기반 수업 모형을 개발하는 것을 목표로 한다. 이를 위해 설계 개발 연구 방법론을 적용하여 모형을 개발하고 타당화 하고자 한다. 설계 개발 연구 방법은 교수적 또는 비 교수적 산출물 및 도구, 그리고 그 개발을 이끄는 새로운 모형의 생성에 관한 실증적 기반 확립을 목표로 하는 설계, 개발, 평가과정에 대한 체계적 연구로, 새로운 지식을 생성할 뿐만 아니라 기존 실천들을 타당화하고 교수설계의 이론기반을 실증적으로 확대하고 발전시키는 데 그 목적을 두고 있다[35]. 설계 개발 연구 방법은 ‘산출물 및 도구 개발 연구’와 ‘모형 연구’로 구분되는데 [9], 본 연구에서는 ‘모형 연구’를 통해 모형을 개발하고 타당화한다.

#### 3.2 연구 절차 및 방법

본 연구의 절차는 Fig. 3과 같다. 우선 모형의 초안과 교수전략 개발 단계에서 선행연구 분석을 통해 데이터 리터러시를 위한 머신러닝 기반 AI 융합 수업 모형의 초안을 개발한다. 그 후, 내적 타당화 과정에서 초기

수업모형에 대한 전문가 검토를 두 차례 실시하고, 전문가의 의견을 바탕으로 수업모형과 교수전략을 수정, 보완한다. 외적 타당화에서는 서울 소재 특성화고등학교 학생 15명을 대상으로 실험반, 대조반을 구성한 후, 학생들을 대상으로 데이터 리터러시 관련 설문지로 사전검사를 실시한다. 이후, 개발된 수업을 실시한 후에 같은 설문지로 데이터 리터러시가 얼마나 향상되었는지를 측정한다.

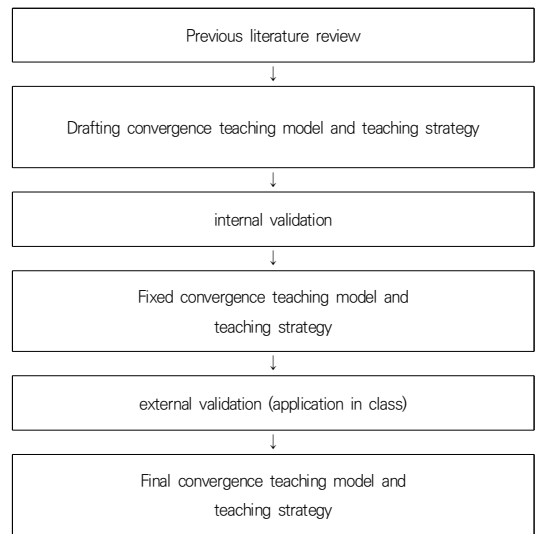


Fig. 3. Research procedures

#### 3.3 연구 참여자

##### 3.3.1 연구 대상

본 연구는 2023학년도 2학기 서울 소재 한 상업계열 특성화고등학교 1학년 학생 총 15명을 대상으로 수업 및 효과검증을 위한 설문 및 면담, 자료수집 등을 하였다. 연구 참여를 위해 학생들과 보호자에게 연구의 개요를 사전 설명하고 동의를 받았다. 해당 학교의 학습자들은 1학년 때 컴퓨터를 사용하는 실무 과목을 수강하기 때문에 기본적으로 컴퓨터 활용 능력을 보유하고 있고, 공통과학 과목을 수강하여 화학에 대한 기초적인 이해도를 갖고 있다. 하지만 1학년에는 정보 관련 교과가 개설되지 않아 데이터를 다루고 분석하는 수업의 경험은 없었다. 또한 사전 조사 결과 프로그래밍 언어에 대한 이해도 거의 전무하여 데이터를 분석하기 위해서 블록 코딩 기반이면서 학생들이 쉽게 접근할 수 있는 엔트리를 사용하였다.

### 3.3.2 전문가 타당화(내적 타당화)

데이터 리터러시를 위한 머신러닝 기반 AI 융합 수업 모형의 타당화를 위하여 전문가 타당화를 실시하였다. 전문가 타당화를 통해 수업모형, 수업 설계 원리, 교수전략의 타당성, 설명력, 유용성, 이해도, 보편성을 평가하였으며, 학생들의 자기 평가 설문 문항인 리터러시 역량 측정 설문 문항의 타당도를 검토하였다. 전문가 타당화에 참여한 전문가들의 프로필은 다음 Table 2와 같다.

Table 2. Profiles of experts

expert	Job	career	degree	major
A	Teacher	6	M.S	computer education
B	Teacher	17	Ph.D.	chemical education
C	Teacher	7	Ph.D. candidate	computer education
D	Teacher	20	M.S	computer education
E	Teacher	2	Ph.D. candidate	chemistry
			M.S	chemical education
F	Teacher	5	M.S	physics education

모든 전문가는 각 분야뿐만 아니라 인공지능 교육에 대해서도 관심을 갖고, 적극적으로 연수에 참여하거나 직접 연수의 강사로 활동하는 전문가를 선정하였다. 또한 전문가 E는 교직 경력은 짧지만, 사교육 경력 등 실제 교육 경력이 풍부하며 화학 분야에서 박사 과정을 진행했기에 적절한 전문가라고 판단하였다.

### 3.4 연구 도구

본 연구에서 사용한 연구 도구는 내적 타당화를 위한 전문가 타당화 검사지, 학습자 대상 수업 반응 설문지, 학습자 대상 면담 설문지, 데이터 분석 수업에 활용한 엔트리(Entry) 등이다.

수업에 앞서 데이터 리터러시를 측정할 수 있는 도구를 개발하기 위해서 선행연구를 검토하였다. 송유경(2021)은 데이터 리터러시를 통계적 분석, 데이터 기반 의사소통의 2개의 영역으로 구분하고, 통계적 분석 안에 데이터 이해, 준비, 분석, 평가의 4가지 영역으로, 데이터 기반 의사소통 안에 데이터 표현, 데이터 기반 의사결정의 2가지 영역으로 구분하고 25개의 하위 문항을 개발하였다[9]. 또한 조연수(2021)은 데이터의 이해,

수집, 해석 및 평가, 관리, 활용의 5가지 영역으로 구분하고 11개의 하위문항을 개발하였다[10]. 본 연구에서는 그 중 알고리즘과 연관성이 있는 7개의 문항을 추출하고 데이터 이해 및 수집, 데이터 분석, 데이터 기반 의사소통의 3가지 영역으로 구분하였다. 또한 다양한 데이터로 머신러닝을 수행할 때 구분의 기준이 되는 특성을 파악할 수 있는 능력을 파악하기 위해 '나는 데이터의 다양한 특성 중 데이터 분석에 필요한 특성들을 추출할 수 있다.'라는 문항을 신규 개발하였다. 검사도구로 활용한 문항은 Table 3과 같다.

Table 3. Data literacy assessment tool

area	question
Data collection and preparation	I understand that there are various types of data[10]
	I can understand the significance of each row and column in the collected data[9]
	When necessary data is unavailable, I can employ alternative methods to obtain the required information[10]
Data analysis and utilization	I can use visualization techniques to analyze data[9]
	I can prepare for data analysis and use appropriate tools to conduct the analysis[10]
	I can extract the necessary features for data analysis from the various characteristics of the data
Data-driven communication	I can explain the results of data analysis in a meaningful way[10]
	I accurately cite and attribute the source when using data[10]

## 4. 연구 결과

### 4.1 초기 수업 모형 및 수업 설계 원리, 상세 지침

선행 문헌을 고찰한 결과 데이터 리터러시를 위한 탐구 단계를 1차로 '문제 정의, 데이터 수집, 데이터 분석, 적용 및 평가'의 순으로 도출하였다[9, 21, 36, 37]. 데이터 중심 인공지능 개념에 의하면 데이터는 전처리 단계에서만 다루어지는 정적인 입력이 아니라, 인간 중심적이고 동적이며 항상 진화하는 것으로 간주되어야 한다. 따라서 데이터 개선은 전처리 단계를 넘어 과정 전반에 걸쳐 중요한 반복 작업이 되어야 한다[13]. 이를

반영하여 머신러닝 실시 이후 데이터 수집 단계가 반복 되도록 하였다.

데이터를 이용한 탐구 활동에서 학생들이 일부라도 데이터 수집자의 역할을 수행하도록 하면 데이터에 대한 비판적 사고능력 향상에 도움이 된다[38]. 현실적 수업 상황을 고려하여 첫 번째 머신러닝 수행 단계에서는 교사가 제공하는 정제된 소규모 데이터를 사용하고 두 번째 머신러닝 수행 단계에서 학생들이 직접 데이터를 추가하는 과정을 거치도록 하였다. 또한 수업 설계를 위한 원리 및 지침을 도출하기 위해 선행 문헌을 고찰하여 수업 모형의 단계에 따라 항목을 구성하였다. 수업을 설계하기 위한 원리는 Table 4와 같다. 그리고 수업 설계 원리를 바탕으로 Table 5의 상세 지침을 개발하였다.

**Table 4. Step and principle of AI convergence model**

step	Lesson Guides	exploring data	machine learning
principle	1. intrinsic motivation 2. tool exploration 3. scaffolding	1. active exploration 2. intentional arrangement	1. understanding principles 2. coaching
step	Analyzing & evaluating	collecting data	representation & feedback
principle	1. communication and collaboration 2. problem specification	participation in data collection	expression and sharing

**Table 5. Detailed instructions for AI convergence model**

principle	detailed instructions
intrinsic motivation	Presents real-world problems through interesting and compelling visual representations[39]
	Rather than having learners solve problems as an application of the learning content, it guides them to learn key content in order to solve problems[39]
tool exploration	Suggests tools that produce data in familiar and simple formats[40]
	Provides guidance on the basic usage of tools and encourages students to independently learn, when necessary, through internet searches, user manuals, and other resources[40]
scaffolding	Facilitates learning the tool's usage through demonstrations using sample data
	Provides students with separate instructional materials on tool usage[41]

active exploration	Explains the data collection process and ensures understanding of the preprocessing steps
	Allows students to explore a subset of data first, understand its relevance to the problem, and then examine the overall data structure[38]
intentional arrangement	Provides organized data after machine learning execution, allowing students to raise questions about the outcomes
understanding principles	Emphasizes competency development rather than just practical skills[38]
	Observes differences in results based on machine learning models, enabling an understanding of the importance of selecting models that align with data characteristics.
coaching	Provides hints or assistance when learners encounter difficulties in their tasks
communication and collaboration	Designs conversational and collaborative devices to facilitate interaction, actively encouraging group activities[42]
problem specification	Encourages learners to articulate specific curiosities arising during the data interpretation process[40]
	Encourages questioning of analysis results and engages in discussions about potential solutions
participation in data collection	Encourages learners to develop critical thinking skills about data by directly engaging in data collection[38]
	Facilitates learners to experience selecting and refining data relevant to the problem they want to solve
	Ensures the collection of reliable data
expression and sharing	Supports learners in creating structured reflective journals and open portfolios for reviewing their overall learning activities[43]
	Utilize tools that provide a predefined framework for expressive means, offering simplicity in usage and the capability to include various file formats[40]
	Provides feedback on the process and results of activities, including self-assessment, peer assessment, and teacher assessment

#### 4.2 전문가 타당화 자료 분석

전문가 타당화를 통해 해당 수업모형, 수업 설계 원리, 교수 전략, 데이터 리터러시 역량 측정 질문의 타당화를 위하여 두 차례의 타당화를 실시하였다. 1, 2차 타당화에서 각 영역별로 4점 척도(4: 매우 타당하다, 3: 타당하다, 2: 타당하지 않다, 1: 전혀 타당하지 않다)로 평가할 것을 요청하였다.

본 연구에서 수집한 전문가의 응답에 대한 타당도와 신뢰도를 검증하기 위하여 내용 타당도 지수(CVI)와 평가자 간 일치도(IRA)를 분석하였다. Lawshe(1975)가

제시한 내용 타당도 지수(CVI)는 각 전문가의 응답 중 긍정 응답(3: 타당하다, 4: 매우 타당하다)의 비율을 계산하여 타당도를 검증하는 방법이며, 각 문항에 대한 CVI 값이 .80 이상일 때 타당하다고 검증할 수 있다 [44, 45]. 평가자 간 일치도(IRA)는 전문가들의 응답이 일치하는 문항의 수를 전체 문항 수로 나눈 것으로 IRA 값이 .80을 초과할 때 신뢰할 수 있는 것으로 간주한다 [46]. 1차 전문가 타당화에서 모형의 타당도에 대한 응답을 분석한 결과는 Table 6, 수업 설계 원리 및 교수 전략에 대한 결과는 Table 7과 같다.

Table 6. Validity of model (1st)

ability	A	B	C	D	E	F	avg	CVI	IRA
validity	3	4	4	3	3	4	3.50	1	0.6
persuasive power	3	4	3	3	3	3	3.17	1	
usefulness	3	4	3	4	3	4	3.50	1	
understanding	3	4	4	3	2	3	3.17	0.83	
universality	4	3	4	3	2	4	3.33	0.83	

Table 7 Validity of the Principles of lesson design and teaching strategy (1st)

ability	A	B	C	D	E	F	avg	CVI	IRA
validity	3	4	3	3	3	4	3.33	1	0.6
persuasive power	3	4	3	3	3	4	3.33	1	
usefulness	3	4	4	4	3	3	3.50	1	
understanding	4	3	4	3	2	4	3.33	0.83	
universality	3	3	4	3	2	4	3.17	0.83	

1차 전문가 타당도 결과 수업 모형, 수업설계 전략 및 교수전략의 내용 타당도는 검증되었으나, 전문가 응답의 신뢰도를 확보하지 못하였다. 또한 데이터 리터러시 역량 측정 도구에서는 6번 문항과 7번 문항에서 질문의 이해도가 다소 떨어진다는 의견을 받았다.

전문가 타당도로 받은 피드백을 반영하여 수업설계 원리 및 교수전략과 데이터 리터러시 역량 측정 도구를 수정하였다. 수업 설계 원리 및 교수전략에서는 수업 안내 부분에서 '원리 이해'를 추가하였으며, 데이터 탐색 부분에서는 '능동적 탐색'의 세부 교수전략을 수정하였다. 머신러닝 수행 부분에서는 '원리 이해'를 수업 모형에 맞게 '역량 개발'로 수정하였으며, 데이터 수집 부분의 단계를 좀 더 구체화하여 수업을 설계하고 교수전략을 구상할 수 있도록 하였다. 또한 데이터 리터러시

역량 측정 도구에서는 5번 문항의 질문을 명료화하였으며, 수업 모형과 관련이 적은 7번 문항을 삭제하였다. 2차 전문가 타당화에서 모형의 타당도에 대한 응답을 분석한 결과는 Table 8, 수업 설계 원리 및 교수 전략에 대한 결과는 Table 9과 같다.

Table 8. Validity of model (2nd)

ability	A	B	C	D	E	F	avg	CVI	IRA
validity	3	4	4	4	4	4	3.83	1	1
persuasive power	3	4	4	4	4	4	3.83	1	
usefulness	3	4	4	4	4	4	3.83	1	
understanding	3	4	4	4	4	4	3.83	1	
universality	3	4	4	4	4	4	3.83	1	

Table 9. Validity of the Principles of lesson design and teaching strategy (2nd)

ability	A	B	C	D	E	F	avg	CVI	IRA
validity	3	4	4	4	4	4	3.83	1	1
persuasive power	3	4	3	4	4	4	3.67	1	
usefulness	4	4	4	4	4	4	4	1	
understanding	4	4	4	4	4	4	4	1	
universality	3	4	4	4	4	4	3.83	1	

2차 전문가 타당화 결과 수업설계 원리 및 교수전략 타당도의 설명력을 제외한 모든 문항에서 내용 타당도와 전문가 응답의 신뢰도를 확보하였다. 2차 전문가 타당도의 피드백을 바탕으로 최종 수업모형, 수업 설계 원리 및 교수전략, 데이터 리터러시 역량 측정 도구를 개발하였다. 최종 개발된 수업모형은 Fig. 4와 같다. 그리고 이에 대한 수정된 최종 수업 설계 원리 및 상세 지침은 Table 10과 같다.

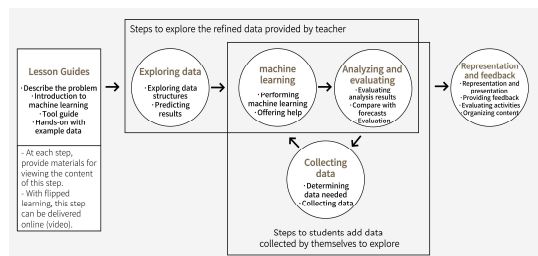


Fig. 4. Fixed model for AI convergence learning



Table 10. Fixed detailed instructions for AI convergence model

principle	detailed instructions
intrinsic motivation	Presents real-world problems through interesting and compelling visual representations[39]
	Rather than having learners solve problems as an application of the learning content, it guides them to learn key content in order to solve problems[39]
tool exploration	Suggests tools that produce data in familiar and simple formats.[40]
	Provides guidance on the basic usage of tools and encourages students to independently learn, when necessary, through internet searches, user manuals, and other resources[40]
scaffolding	Facilitates learning the tool's usage through demonstrations using sample data
	Provides students with separate instructional materials on tool usage[41]
understanding principles	Emphasizes competency development rather than just practical skills[38]
	Observes differences in results based on machine learning models, enabling an understanding of the importance of selecting models that align with data characteristics.
active exploration	Explains the data collection process and ensures understanding of the preprocessing steps
	Allows students to explore a subset of data first, understand its relevance to the problem, and then examine the overall data structure[38]
intentional arrangement	Provides organized data after machine learning execution, allowing students to raise questions about the outcomes
principle	detailed instructions
coaching	Provides hints or assistance when learners encounter difficulties in their tasks
communication and collaboration	Designs conversational and collaborative devices to facilitate interaction, actively encouraging group activities[42]
problem specification	Encourages learners to articulate specific curiosities arising during the data interpretation process[40]
	Encourages questioning of analysis results and engages in discussions about potential solutions
participation in data collection	Encourages learners to develop critical thinking skills about data by directly engaging in data collection[38]
	Facilitates learners to experience selecting and refining data relevant to the problem they want to solve
	Ensures the collection of reliable data
understanding principles	Adjust parameters and observe changes in the model's accuracy
knowledge integration	Enable the interpretation of machine learning results in conjunction with disciplinary knowledge[40]
expression and sharing	Supports learners in creating structured reflective journals and open portfolios for reviewing their overall learning activities[43]
	Utilize tools that provide a predefined framework for expressive means, offering simplicity in usage and the capability to include various file formats[40]
	Provides feedback on the process and results of activities, including self-assessment, peer assessment, and teacher assessment

### 4.3 외적 타당화 결과

#### 4.3.1 외적 타당화 개요

본 연구에서는 데이터 리터러시를 향상시킬 수 있는 머신러닝 기반 융합 수업 모형을 개발하는 것이고, 이를 현장에 적용함으로써 개선 사항을 반영하여 최종 모형과 교수학습 전략을 개발하는 것이다. 이를 위해 2023년 11월부터 12월까지 서울 소재 특성화고등학교 1학년 학생 15명을 대상으로 3차시 분량의 수업을 실시하였다. 사전에 학생들의 데이터 리터러시를 판단하기 위해 데이터 리터러시 역량을 검사하였으며, 학생들의 컴퓨터언어 사용 수준 또한 함께 조사하였다. 대상 학생 15명 모두 엔트리(Entry)에 대해서는 사용 경험이 있었으나, 파이썬을 접해본 학생은 1명에 불과하여 수업 진행간 학생들에게 친숙한 엔트리를 사용하였다. 수업이 완료된 후에는 데이터 리터러시 검사를 다시 실시하여 데이터 리터러시의 향상 정도를 측정하였고, 각 차시가 종료될 때마다 활동지를 통해 학생들의 이해 수준을 평가하였다.

수업 절차는 본 연구에서 개발한 수업 모형에 따르며, 내용은 고등학교 1학년 통합과학에서 다루는 '물질의 규칙성'을 탐구 주제로 설정하였다. 물질의 끓는점과 녹는 점은 1학년 학생들이 이미 숙지하고 있는 내용으로, 해당 특성들로 여러 화합물을 결정트리를 활용하여 분류하는 것을 목표로 하였다. 수업 차시는 총 3차시 분량으로 구성되었으며, 각 차시의 구성과 본 수업 단계는 Table 11과 같다. 매 차시마다 활동지를 제작하여 학생들이 산출물을 제작하도록 했으며, 학생들의 성취 수준을 확인할 수 있도록 하였다. 수업 환경은 학생들이 개인별로 컴퓨터를 사용할 수 있는 실습실에서 진행하였다.

Table 11. Curriculum of AI convergence Education based on machine learning

Session	Steps	Contents
1	Lesson Guides	- Understanding Machine Learning 1. Learning about Entry and guiding how to use Entry to utilize artificial intelligence models 2. Creating a decision tree with basic example data 3. Enhance understanding of the principles of decision trees by learning with diverse datasets
		- Exploring chemical data
2	Exploring data	Exploration of compound data, including exploration of the meanings of rows and columns in the dataset
	Machine learning	Creating a decision tree using Entry Enhancing the accuracy of the decision tree by modifying learning conditions

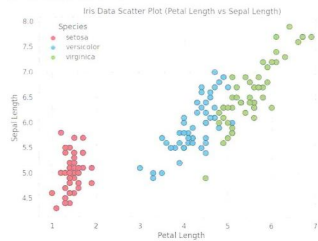
(Continued)

Table 11. Curriculum of AI convergence Education based on machine learning

Session	Steps	Contents
2	Analyzing and evaluating	Analyzing and evaluating a decision tree's accuracy
	collecting data	Collecting additional data using generative AI
3	Machine learning (2nd)	Creating a decision tree using Entry(2nd)
	Analyzing and evaluating (2nd)	Analyzing and evaluating a decision tree's accuracy
	representation & feedback	presentation and feedback about decision tree

1차시 수업은 수업 안내(Lesson Guides) 단계로서 결정트리를 이해하는 데 목표를 두고 구성하였다. 내재적 동기유발(intrinsic motivation)의 원리에 따라 머신러닝의 실제 사용 사례를 들었고, 도구 탐색(tool exploration)의 원리에 따라 결정트리의 기본 사용법을 알려주고, 설명서 활동지를 통해 스스로 익힐 수 있도록 지도하였다. 스캐폴딩(scaffolding)의 원리에 따라 간단한 결정트리를 통해 데이터가 산출되는 것을 보여주었고, 이를 통해 결정트리의 중요성을 이해하여 원리 이해(understanding principles)원리가 이루어지도록 하였다. 끝으로 엔트리에서 제공하는 붓꽃 예시 데이터를 통하여 학생들이 결정트리를 작성하도록 지도하였다. 1차시 수업 후 학생들이 산출한 활동지의 예시는 Fig. 5와 같으며, 기본적인 결정트리에 대해 이해한 것을 볼 수 있다.

#2. 다음은 붓꽃(Iris)의 3가지 종(Setosa, Versicolor, Virginica)을 꽃잎의 길이(Petal Length)와 꽃받침 길이(Sepal Length)로 구분한 그림이다. 다음 그림을 보고, 각 종을 구분할 수 있는 기준이 있는지 알아서 보자.



#3. 엔트리를 활용하여 붓꽃 데이터를 분류할 수 있는 결정트리를 그려보시오.

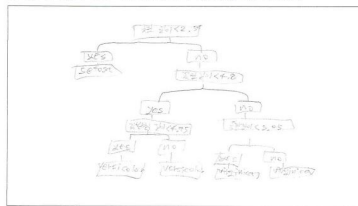


Fig. 5. Decision trees produced by students(1st session)

2차시 수업은 데이터 탐색(Exploring data), 머신러닝(Machine learning), 분석 및 평가(Analyzing and evaluating) 단계로, 본 연구의 주 데이터인 화합물 데이터에 대해 분석, 결정트리를 작성하는 것을 목표로 구성하였다. 우선 엔트리에서 제공하는 타이타닉 생존 데이터, 펭귄 종별 데이터, 티셔츠 신체 지수 데이터로 결정트리를 작성하며 결정트리 제작을 연습하여 결정트리에 대한 이해도를 높였다. 이후 학생들에게 사전에 전처리된 화합물들의 끓는점, 녹는점 데이터를 제공하였고, 능동적 탐색(active exploration)의 원리에 따라 화합물 데이터를 시각화하며 데이터를 탐색하였다. 다음으로 역량 개발(intentional arrangement)과 코칭(coaching)의 원리에 따라 학생들이 데이터 리터러시를 함양하면서 머신러닝을 체험할 수 있도록 도움을 제공하였다. 끝으로 제작한 화합물 결정트리를 소통과 협력(communication and collaboration) 원리에 따라 학생들끼리 상호작용하면서 서로의 정확도, 특성 등에 대해 의사소통 하도록 독려하고, 문제 구체화(problem specification)의 원리에 따라 정확도 개선이나 분류 기준 등에 대해 개선 방안을 탐구하도록 하였다. 2차시 수업 진행 간 학생들이 산출한 활동지는 Fig. 6과 같다.

오늘의 QUEST!  
 새로운 데이터를 결정트리로 만들어보기!  
 - 결정트리를 사용하기 위해서는 분류가 되어있는 데이터를 사용해야 합니다 !!  
 예시) 엔트리 데이터 중  
 1. 타이타닉 예시 데이터(생존/사망 분류)  
 2. 티셔츠 사이즈 예시 데이터(사이즈별 분류)  
 3. 펭귄 예시 데이터(종별 분류)  
 4. 유산병의 화합물 데이터(약합물별 분류)

1. 자신이 사용한 데이터 :
2. 분류 기준으로 삼은 특성 :
3. 자신이 만든 결정트리를 아래에 그려봅시다.

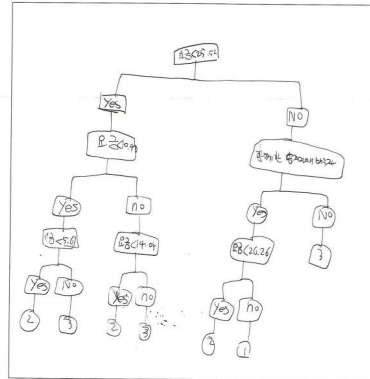


Fig. 6. Decision trees produced by students(2nd session)

3차시 수업은 데이터 수집(Collecting data), 2차 머신러닝(Machine learning), 분석 및 평가(Analyzing and evaluating), 표현 및 피드백(Representation and feedback)의 단계로 새롭게 수집한 데이터를 바탕으로 화합물 결정트리를 보완, 작성하는 것을 목표로 하였다. 2차시에 작성했던 결정트리를 바탕으로 부족한 데이터를 수집하였다. 데이터 수집 방법으로는 ChatGPT와 Bing AI 등 생성형 AI를 활용하여 검색하였고, 새롭게 수집한 데이터를 기존의 데이터 셋에 더한 후 새롭게 결정트리를 제작하였다. 학생들은 제작한 데이터를 표현과 공유(expression and sharing)원리에 따라 발표하고 공유하였으며, 서로의 결정트리를 비교해보았다. 학생들은 이전에 비해 보다 정교한 결정트리를 원활하게 작성하였으며, 결정트리에 대한 이해도가 증가하여 노드의 최대 깊이, 노드의 최소 데이터 수 등 다양한 파라미터등을 조절하며 정확도를 개선하려는 노력 또한 보여주었다. 3차시 수업 진행 간 학생들이 산출한 활동지는 Fig. 7과 같다.

시키는 데 있다. 이를 위해 본 연구에서는 데이터 리터러시 검사 도구를 활용하여 단일집단 사전, 사후 검사를 실시하였다. 개발한 데이터 리터러시 검사도구의 신뢰성을 검증하기 위해 학습자 15명을 대상으로 실시한 사전 검사 결과에 대해 크론바흐 알파(Cronbach's  $\alpha$ ) 검사를 실시하였다. 그 결과 전체 0.946으로 높은 수준의 신뢰도를 보였다. 크론바흐 알파 검사의 결과를 표로 나타내면 Table 12와 같다.

Table 12. Cronbach's alpha test results of data literacy Assessment tool (n=15)

No.	Pre-test avg	Pre-test std	cronbach alpha
1	2.87	1.19	0.946
2	2.40	1.18	
3	2.13	0.92	
4	2.27	0.96	
5	2.33	0.98	
6	2.20	0.94	
7	2.20	0.86	
8	2.53	1.19	
	2.37	1.03	

서울대학교목표과 알고리즘 수업 활동지

오늘의 QUEST!  
새로운 데이터를 포함시켜 학습시키기

1. 자신이 추가한 데이터 (원하는 개수만큼 더하기)

물질 이름	녹는점	끓는점	분자량	Label
1 hydride	800	1500		hydride
2 ether	-116.3	39.6		ether
3 Carboxylic acid	16.17	118.1		Carboxylic acid
4 aromatic	9.5	80.1		aromatic
5				
6				
7				
8				
9				
10				

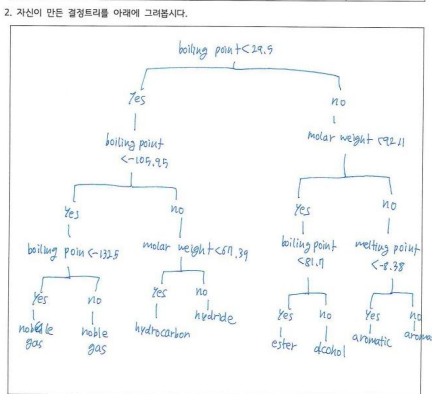


Fig. 7. Decision trees produced by students(3rd session)

4.3.2 수업 사전-사후 데이터 리터러시 향상 여부  
본 수업의 목적은 학습자의 데이터 리터러시를 향상

사전검사의 전체 평균은 2.37로 다소 낮게 나왔는데, 연구 대상반의 경우 특성화고등학교 1학년 학생로서, 데이터 관련 교과가 개설되지 않아 데이터 리터러시가 다소 떨어져 있던 것으로 판단된다.

데이터 리터러시 사전-사후 대응표본 t-검정 분석 결과는 Table 13과 같다. 사후검사의 전체 평균은 4.21로 사전검사에 비해 데이터 리터러시 역량의 유의미한 향상이 있었다.

Table 13. Paired t-test results of pre-post Data literacy (n=15)

No.	pre avg	pre std	post avg	post std	t	p
1	2.87	1.19	4.27	0.59	-4.17	< 0.001***
2	2.40	1.18	4.20	0.56	-5.08	< 0.001***
3	2.13	0.92	4.07	0.70	-5.85	< 0.001***
4	2.27	0.96	4.27	0.88	-5.68	< 0.001***
5	2.33	0.98	4.13	0.74	-5.28	< 0.001***
6	2.20	0.94	4.20	0.77	-6.48	< 0.001***
7	2.20	0.86	4.27	0.80	-6.88	< 0.001***
8	2.53	1.19	4.27	0.96	-4.67	< 0.001***
Sum	2.37	1.03	4.21	0.74		

## 4.3.3 학습자 반응 분석

3차시에 걸친 수업 이후 데이터 리터러시 사후 검사와 함께 학습자 반응을 설문지를 통해 수집하였다. 질문은 모두 서술형으로 제출되었으며 수업에 관해 좋았던 점, 아쉬웠던 점, 더 알고 싶은 점으로 구분하여 서면으로 제출받았다. 먼저, 수업과 관련하여 좋았던 점에 대한 학습자의 반응은 다음과 같다. 대부분 엔트리로 결정트리를 구현하는 것에 대해 흥미를 느꼈으며, 결정 트리에 유용성에 대해 이해하였다.

- (학생 A) 초등학교, 중학교 때는 엔트리로 간단한 코딩만 했었는데 결정트리라는 것을 처음 배워서 신기하고 좋았어요.
- (학생 B) 결정트리가 데이터를 분류한 것을 한눈에 보기 쉽게 시각화 한 것이 좋았어요.
- (학생 C) 데이터를 정리해 새롭게 분류하는 방법에 대해 알게되어 좋았어요.

반면 학생들에게 다소 생소한 결정트리를 사용한 것을 어려웠던 점으로 꼽았다. 대부분의 학생들이 머신러닝을 처음 접하다보니 인공지능이 학습하는 것을 이해하는데 어려움을 겪기도 하였다.

- (학생 D) 결정트리가 처음에는 다소 생소해서 어려웠어요.
- (학생 E) 엔트리로 결정트리를 작성하는 것을 처음해 보니 처음에는 다소 어려웠어요.
- (학생 F) 학습 조건을 어떻게 설정해야 더 정교한 결정트리가 나오는지 이해하기 어려웠어요.

한편 결정트리를 학습하면서 오늘 배운 결정트리를 다양한 분야에서 활용하고 싶어 하거나 결정트리의 정확도를 높일 수 있는 방법에 대해 궁금해하는 학생도 있었다. 본 수업 이후 더 배우고 싶은 내용들은 다음과 같았다.

- (학생 G) 결정트리가 어떤 원리로 데이터를 분류하고 생성하는지 궁금해요.
- (학생 H) 엔트리로 만들 수 있는 다른 머신러닝이 궁금해요.
- (학생 I) 엔트리가 아닌 다른 프로그래밍 언어로 결정트리를 구현할 수 있는지 궁금해요.

## 5. 논의 및 결론

본 연구의 목적은 머신러닝을 기반으로 한 AI 융합 수업 모형과 교수전략을 개발하는 데 있다. 이를 위해 고등학교 과학과 교과 내용을 바탕으로 데이터 리터러시 역량을 향상시키기는 수업모형 및 교수전략을 개발하여 전문가 검토를 통한 내적 타당화와 서울 소재 특성화고등학교 1학년 학생을 대상으로 실제 수업에 적용하여 모형의 타당성을 확보하고 최종적인 모형을 도출하였다. 최종 수업 모형을 적용한 결과 학생들의 데이터 리터러시 역량이 전체 평균 사전 2.37점에서 사후 4.21점으로 향상되어 본 수업 모형이 학생들의 데이터 리터러시 역량 향상에 유의미한 영향을 주었음을 알 수 있었다. 또한 학생들의 학습자 반응 설문 답변을 통해 수업 모형에 대한 다양한 의견을 들을 수 있었다.

본 연구의 의의는 다음과 같다. 첫째, 일반교과와 머신러닝을 융합한 AI융합 교육 수업 모형을 개발하였다. AI를 일반교과와 연결하여 실제 수업으로 적용한 사례는 많지 않았는데, 본 연구를 통해 머신러닝이라는 AI 요소를 일반교과와 융합한 수업모형과 교수전략을 개발하였다.

둘째, 머신러닝과 데이터 리터러시에 관한 상관관계에 대해 연구하였다. 머신러닝은 데이터를 기반으로 이루어지지만, 실제로 머신러닝과 관련된 수업이 데이터 리터러시 향상에 영향을 주는지에 관한 연구는 많지 않았다. 본 연구는 머신러닝을 활용한 수업을 통해 학생들의 데이터 리터러시 역량이 향상됨을 이끌어내어 머신러닝이 데이터 리터러시와 직접적인 연관이 있음을 보여주었다.

셋째, 직업계 고등학교에서 인공지능을 융합한 수업의 효과성을 보여주었다. 본 연구에서 실시한 학생들의 데이터 리터러시 검사에 의하면 학생들의 데이터 리터러시 역량이 전체 평균 사전 2.37점으로 다소 낮게 측정되었다. 상대적으로 데이터와 관련된 교육과정이 부족한 직업계고등학교에서는 학생들의 기본적인 데이터 리터러시 역량이 다소 부족함을 알 수 있었고, 이를 함양하기 위해서는 인공지능을 융합한 다양한 수업모형이 개발되어 직업계 고등학교에서 적용되어야 할 것이다.

본 연구의 제한점은 다음과 같다. 첫째, 직업계 고등학교 학생들의 다소 부족한 프로그래밍 역량으로 파이썬 등의 텍스트 기반 프로그래밍 언어를 사용하지 못하고 엔트리라는 블록코딩을 사용하여 머신러닝을 제한

적으로 사용하였다. 머신러닝의 기능을 온전하게 활용하려면 파이썬 등을 활용하여 텍스트 코딩하는 것이 필요하다. 하지만 연구 대상이었던 직업계 고등학교의 경우 파이썬 언어를 학습하는 교과가 개설되지 않아 부득이 학생들에게 익숙한 엔트리를 사용하게 되었고, 결정 트리의 파라미터를 제한적으로 조작하게 되었다.

둘째, 3차시라는 부족한 수업 차시로 인해 인공지능 융합교육 구성에 다소 아쉬움이 있었다. 다수의 전문교과가 개설된 직업계 고등학교 특성상 많은 시간을 연구에 할애할 수 없어서 부득이 3차시에 걸쳐 수업을 구성하였고, 이로 인해 머신러닝의 요소를 제한적으로 경험하였다.

셋째, 단일 집단으로 진행된 연구로 대조군의 부재하여 수업 모형의 효과성을 입증하는데 제한점이 있었다. 또한 질병, 수업 결손 등의 이유로 도중 일부 학생들이 이탈하여 적용 표본이 다소 부족하였다. 보통 교과 수업 시수가 부족한 직업계 고등학교 특성상 여러 학급에서 융합 수업을 실시하는데 어려움이 있었고, 이에 따라 대조군 및 많은 수의 설문 표본을 확보하지 못하였다.

본 연구의 기대효과 및 제언은 다음과 같다. 첫째, 다양한 학교급과 교과에서 본 연구의 수업모형을 기반으로 한 머신러닝 융합 수업이 이루어질 수 있다. 다양한 교과에서 데이터가 발생하는 지점에서 머신러닝 기반 융합수업을 구성함으로써 학생들의 데이터 리터러시를 향상시키고 머신러닝의 유용성을 인식시킬 수 있다. 둘째, 화합물이라는 실제 데이터를 활용하여 수업한 사례를 통해, 실생활에 적용 가능한 다양한 데이터 수업 사례를 제시하였다. 마지막으로 화학 교과뿐만 아니라 다양한 일반 교과에서 활용할 수 있는 데이터를 확보하면 향후 머신러닝 기반 융합수업을 진행함에 있어서 도움이 될 것이다. 이를 기반으로 다양한 교과에서 머신러닝을 융합한 수업에 대한 연구가 이루어지길 기대한다.

## REFERENCES

- [1] Kim, S. H. (2019). Analysis of International Educational Trends and Learning Tools for Artificial Intelligence Education. *The Journal of Korean Association of Computer Education*, 23(2), 25-28.
- [2] Lin, P., & Brummelen, J. V. (2021). Engaging Teachers to Co-Design integrated AI Curriculum for K-12 Classrooms. *CHI '21: PROCEEDINGS OF THE 2021 CHI CONFERENCE ON HUMAN FACTORS IN COMPUTING SYSTEMS*. NEW YORK: Assoc Computing Machinery. DOI : 10.1145/3411764.3445377
- [3] The Government of the Republic of Korea. (2020). *Education policy direction and key tasks in the artificial intelligence era*.
- [4] Ministry of Education. (2022). *The 2022 revised curriculum*.
- [5] Lee, D. Y., & Han, S. K. (2022). Analysis of changes in AI competency, attitude, and perception through development and application of AI education program. *Journal of the Korean Association of Artificial Intelligence Education*, 3(3), 7-14. DOI : 10.52618/aied.2022.3.3.2
- [6] Kim, S. M., & Park, K. B. (2023). The effectiveness of K-means clustering algorithm-based class in elementary social education. *Korean Association For Learner-Centered Curriculum And Instruction*, 23(5), 127-140. DOI : 10.22251/jlcci.2023.23.5.127
- [7] Hwang, H. S. (2016). Study on Big Data Utilization in Social Studies Education. *Social Studies Education*, 55(3), 75-89.
- [8] Wolff, A., Gooch, D., & Kortuem, G. (2016). *Data Literacy to Support Human-centred Machine Learning*. In: CHI 2016, 7-12 May 2016, San Jose California, USA. [http://www.doc.gold.ac.uk/~mas02mg/HCML2016/HCML2016\\_paper\\_1.pdf](http://www.doc.gold.ac.uk/~mas02mg/HCML2016/HCML2016_paper_1.pdf)
- [9] Song, Y. K. (2021). *The Data-Driven Debate (DDD) Instructional Model for Improving Data Literacy*. Master's thesis. Seoul National University, Seoul.
- [10] Cho, Y. S. (2022). *Effects of AI Convergence Science Classes on Promoting Middle School Students' Attitude Towards AI Technology and Data Literacy*. Master's thesis. Ewha Womans University, Seoul.
- [11] Jeong, J. H., Kim, J. Y., & Kim, K. H. (2022). A Research on the Design and Application of a Machine Learning Project Class Programs in Artificial Intelligence Education in High School. *The Journal of Korean Association of Computer Education*. 26(1), 203-204.
- [12] Olari, V., & Romeike, R. (2021). Addressing AI and Data Literacy in Teacher Education: A Review of Existing Educational Frameworks. In *Proceedings of the 16th Workshop in Primary*

- and Secondary Computing Education (WiPSCE '21). *Association for Computing Machinery*, New York, NY, USA, Article 17, 1-2.  
DOI : 10.1145/3481312.3481351
- [13] Jarrahi, M. H., Memariani, A., & Guha, S. (2022). The Principles of Data-Centric AI (DCAI). *Communications of the ACM*, 66(8), 84-92.  
DOI : 10.1145/3571724
- [14] Strickland, E. (2022). Andrew Ng, AI Minimalist: The Machine-Learning Pioneer Says Small is the New Big. In *IEEE Spectrum*, 59(4), 22-50.  
DOI : 10.1109/MSPEC.2022.9754503
- [15] AI4K12 (2021). *Grade Band Progression Charts*. <https://ai4k12.org/gradeband-progression-charts/>
- [16] Actua (2022). *Actua's Artificial Intelligence (AI) Education Handbook*. CIRA. [https://actua.ca/wp-content/uploads/2023/07/AI\\_Handbook\\_2023\\_V2.pdf](https://actua.ca/wp-content/uploads/2023/07/AI_Handbook_2023_V2.pdf)
- [17] Long, D., & Magerko, B. (2020). What is AI Literacy? Competencies and Design Considerations. *Conference on Human Factors in Computing Systems - Proceedings*. NEW YORK: Assoc Computing Machinery.  
DOI : 10.1145/3313831.3376727
- [18] Wolff, A., Gooch, D., Cavero Montaner, J. J., Rashid, U., & Kortuem, G. (2016). Creating an Understanding of Data Literacy for a Data-driven Society. *The Journal of Community Informatics*, 12(3), 9-26.  
DOI : 10.15353/joci.v12i3.3275
- [19] Lee, W. T. (2015). Overcoming Information Disparities in the Era of Big Data through Data Literacy. *KISO JOURNAL*, 21. <https://journal.kiso.or.kr/?p=7012%202015.12.21>
- [20] Han, S. W. (2018). A Study about the Concept of Data Literacy based on Digital Humanities. *Journal of the Korean Society for Information Management*, 35(4), 223-236.  
DOI : 10.3743/KOSIM.2018.35.4.223
- [21] Bae, H. S. (2019). Educational Implications of Data Literacy in Social Studies. *Theory and Research in Citizenship Education*, 51(1), 95-120.  
DOI : 10.35557/trce.51.1.201903.004
- [22] Lee, E. J. (2022). *The Effects of AI-based Data Analysis Education on Convergent Thinking Ability and Data Literacy of General High School Students*. Master's thesis. Kongju National University, Gongju.
- [23] Korea Institute for Health and Social Affairs, Oh, M. A., Choi, H. S., Kim, S. H., Chang, J. H., Jin, J. H. & Cheon, M. K. (2017). *A Study on Social security Big Data Analysis and Prediction Model based on Machine Learning*. (2017-46). <https://www.kihasa.re.kr/publish/report/view?seq=27848>
- [24] Cho, M. H. (2021). A Study on the History, Classification and Development Direction of Artificial Intelligence. *Journal of The Korea Institute of Electronic Communication Sciences*, 16(2), 307-312.  
DOI : 10.13067/JKIECS.2021.16.2.307
- [25] Mariescu-Istodor, R., & Jormanainen, I. (2019). Machine Learning for High School Students. *19TH KOLI CALLING CONFERENCE ON COMPUTING EDUCATION RESEARCH (KOLI CALLING 2019)*. NEW YORK: Assoc Computing Machinery.  
DOI : 10.1145/3364510.3364520
- [26] Kim, D. Y., Chae, D. Y., & Park, S. H. (2023). Development and application of PBL-based machine learning education program to improve elementary school students' problem solving skills. *Korean Association For Learner-Centered Curriculum And Instruction*, 23(6), 639-661.  
DOI : 10.22251/jlcci.2023.23.6.639
- [27] Moon, W. J. et al. (2021). Effect of Machine Learning Education Focused on Data Labeling on Computational Thinking of Elementary School Students. *Journal of The Korean Association of Information Education*, 25(2), 327-335.  
<https://doi.org/10.14352/jkaie.2021.25.2.327>
- [28] *A Study on how to apply AI education to K-12*.(2023). Seoul. Korea Foundation for the advancement of Science & Creativity.
- [29] Korea Institute for Curriculum and Evaluation, Hong, S. J. et al. *Artificial Intelligence and EduTech in School Education*. (RRI 2020-2). <https://www.kice.re.kr/>
- [30] Kwon, H. S. et al. (2021). Current Status of the Implementation of Convergence Education in Primary and Secondary Schools. *Journal of Science Education*, 45(3), 336-348.  
DOI : 10.21796/jse.2021.45.3.336
- [31] Park, J. Y. (2023). Analysis of Attitude Toward AI According to SW Non-major's Computational Thinking and AI Experience. *The Journal of Korean Association of Computer Education*, 26(1), 33-41.  
DOI : 10.32431/kace.2023.26.1.004
- [32] Leem, J. H. (2018). Main Issues of Software

- Education and Tasks of Educational Technology for improving Software Education. *Journal of Educational Technology*, 34(3), 679-709.  
DOI : 10.17232/KSET.34.3.679
- [33] Noh, D. K. (2023). *Development and application of artificial intelligence (AI) and high school science integrated education programs based on scientific data*. Master's thesis. Seoul National University, Seoul.
- [34] Lee, J. H., Jo, J. H., & Chae, S. C (2021). Development of Data-driven Teaching Material for AI Convergence Education: Focused on Damped Oscillation. *School Science Journal*, 15(2), 121-134.
- [35] Lim, C. I. (2012). *Curriculum Design Theories and Models*. Seoul : KYOYOOKBOOK.
- [36] Mandinach, E. B., & Gummer, E. S. (2016). What does it mean for teachers to be data literate: Laying out the skills, knowledge, and dispositions. *Teaching and Teacher Education*, 60, 366-376.  
DOI : 10.1016/j.tate.2016.07.011
- [37] Kim, B. C., Kim, B. S., & Kim, J. H. (2022). Development and Validation of Data Science Education Instructional Model. *Journal of The Korean Association of Information Education*, 26(5), 417-425.  
DOI : 10.14352/jkaie.2022.26.5.417
- [38] Wolff, A., Wermelinger, M., & Petre, M. (2019). Exploring design principles for data literacy activities to support children's inquiries from complex data. *International Journal of Human-Computer Studies*, 129, 41-54.  
DOI : 10.1016/j.ijhcs.2019.03.006
- [39] Jonassen, D. (1999). 10 Designing Constructivist Learning Environments. Instructional-design theories and models, 11, 21.  
<https://www.davidlewisphd.com/courses/EDD8121/readings/1999-Jonassen.pdf>
- [40] Son, M. H., & Jeong, D. H. (2020). Development of Data-Driven Science Inquiry Model and Strategy for Cultivating Knowledge-Information-Processing Competency. *Journal of the Korean Association for Science Education*, 40(6), 657-670.  
DOI : 0.14697/jkase.2020.40.6.657
- [41] Seo, Y. N., Noh, J. Y., Park, M. R., & Jung, S. J. (2023). A Developmental study of an Instructional Model for Inquiry of Social Studies Based on Data-driven Artificial Intelligence Convergence Education. *Korean Association For*
- Learner-Centered Curriculum And Instruction*, 23(12), 1-25.  
DOI : 10.22251/jlcci.2023.23.12.1
- [42] Hong, H. W. (2020). *Development of the Design Principles of Constructivist Learning Environment for Teaching Pre-Service Teachers in Elementary School Teaching Practices*. Doctoral dissertation. Jeonbuk National University, Jeollabuk-do.
- [43] Chee, H. K., & Lim, C. I. (2022). A Study on Development of an Instructional Design Principles of Integrating Subject Matter and Software for Creative Problem Solving. *Journal of Educational Technology*, 38(2), 369-407.  
DOI : 10.17232/KSET.38.2.369
- [44] Lawshe, C. H. (1975). A quantitative approach to content validity. *Personnel Psychology*, 28(4), 563-575.  
DOI : 10.1111/j.1744-6570.1975.tb01393.x
- [45] Grant, J. S., & Davis, L. L. (1997). Selection and use of content experts for instrument development. *Research in nursing & health*, 20(3), 269-274.  
DOI : 10.1002/(SICI)1098-240X(199706)20:3<269::AID-NUR9>3.0.CO;2-G
- [46] Lynn, M. R. (1986). Determination and quantification of content validity. *Nursing research (New York)*, 35(6), 382-386.  
DOI : 10.1097/00006199-198611000-00017

이 유 진(Yoo-Jin Lee)

【정회원】



- 2017년 2월 : 고려대학교 수학교육과 (이학사)
- 2017년 3월 ~ 현재: 대동세무고등학교 교사
- 2022년 8월 ~ 현재 : 서울대학교 AI융합교육학과 석사과정

- 관심분야 : AI융합교육, 머신러닝
- E-Mail : yoo512@snu.ac.kr

강 상 우(Sang-Woo Kang)

【정회원】



- 2015년 2월: 홍익대학교 수학교육과 (이학사)
- 2016년 3월 ~ 현재: 한영고등학교 교사
- 2022년 8월 ~ 현재 : 서울대학교 AI융합교육학과 석사과정

- 관심분야 : AI융합교육, 인공지능수학
- E-Mail : swgang@snu.ac.kr

임 호 정(Hyo-Jeong Lim)

[정회원]



- 2008년 2월 : 서울대학교 화학교육과, 수학교육과(이학사)
- 2009년 3월 ~ 2014년 2월 : 성동공업고등학교 교사
- 2014년 3월 ~ 2019년 2월 : 휘봉고등학교 교사
- 2019년 3월 ~ 현재 : 개포고등학교 교사
- 2022년 8월 ~ 현재 : 서울대학교 AI융합교육학과 석사과정
- 관심분야 : AI융합교육, 수학교육
- E-Mail: sudgarak@gmail.com

최 원 근(Won-Keun Choi)

[정회원]



- 2017년 2월: 서울대학교 화학교육과(이학사)
- 2017년 3월 ~ 현재: 서울과학고등학교 교사
- 2022년 8월 ~ 현재: 서울대학교 AI융합교육학과 석사과정
- 관심분야 : AI융합교육, 머신러닝, 화학교육
- E-Mail : dnjsrms555@gmail.com