

K-NN 알고리즘 이해를 기반한 머신러닝 교육 프로그램의 개발 및 적용

최은선*, 박남제**

제주대학교 일반대학원 과학교육학부 컴퓨터교육학과*, 제주대학교 교육대학 초등컴퓨터교육전공**

요약

제4차 산업혁명 시대의 도래 후 국가 경쟁력의 가장 중요한 요소는 인공지능이라고 할 수 있다. 그러나, 인공지능의 발달은 빅데이터를 기반한 머신러닝 없이는 불가능하기 때문에 미래형 인재는 머신러닝을 어릴 때부터 친숙하게 활용하고 이해하여 실생활에 창의적으로 적용할 수 있는 역량이 필수적으로 수반되어야 한다. 이에 본 논문에서는 K-NN 알고리즘 이해를 기반한 초등학생 대상 머신러닝 교육 프로그램을 개발하고 적용하여 그 효과를 분석하였다. 그 결과 교육 프로그램을 적용한 학생들의 K-NN 알고리즘에 대한 이해를 바탕으로 머신러닝 원리와 적용에 대한 이해도가 향상되었음을 알 수 있었다. 이를 통하여 제안된 교육 프로그램의 효과가 검증되었으며, 정보 기술에 대한 기초 지식 수준이 낮은 학생들에게도 머신러닝을 교육시킬 수 있는 방안에 대한 가능성을 확장시켰다는 관점에서 본 논문의 의의를 찾을 수 있다.

키워드 : 머신러닝 교육, 인공지능, K-NN, 교육 프로그램, 초등교육

Application and Development of Machine Learning Training Program based on Understanding K-NN Algorithm

Eunsun Choi*, Namje Park**

Major in Computer Education, Faculty of Science Edu., Graduate School, Jeju National University*

Department. of Computer Education, Teachers College, Jeju National University**

Abstract

After the advent of the 4th industrial revolution era, the essential element of national competitiveness can be said to be artificial intelligence. However, since artificial intelligence is impossible without machine learning based on big data, future talents must be accompanied by the ability to use and understand machine learning familiarly from childhood and apply creatively to real life. In this paper, we developed and applied a machine learning education program based on understanding the K-NN algorithm for elementary school students and analyzed the effect. As a result, we found that the students' understanding of machine learning principles and machine learning application were improved based on understanding the K-NN algorithm. Through these results, the effectiveness of the proposed education program was verified. The significance of this paper can be found from the viewpoint of expanding the possibility of a way to educate machine learning even for students with low basic knowledge of information technology.

Keywords : Machine Learning Education, Artificial Intelligence, K-NN, Training Program, Elementary Education

이 논문은 2019년 대한민국 교육부와 한국연구재단의 지원을 받아 수행된 연구임(NRF-2019S1A5C2A04083374).

교신저자: 박남제(제주대학교 초등컴퓨터교육전공)

논문투고 : 2021-01-20

논문심사 : 2021-01-20

심사완료 : 2021-02-10

1. 서론

최근 몇 년간 교육계뿐만 아니라 산업계 전반에서 인공지능(Artificial Intelligence, AI)은 가장 중요한 화두로 떠오르고 있다. 더불어, 인공지능을 가능케 하는 머신러닝(Machine Learning, ML), 딥러닝(Deep Learning, DL), 빅데이터(Big data)와 같은 개념에 대한 관심 또한 증폭되었다[4]. 특히, 머신러닝은 인공지능의 핵심 하위 분야이며, 특정 프로그램 없이 기계가 스스로 학습할 수 있는 능력을 제공하는 분야이다[5]. 이에 인공지능 기술의 실현을 위해서는 머신러닝에 대한 인식과 이해도의 제고가 필연적으로 수반되어야 한다[3].

그러나, 머신러닝의 개념을 모든 학습자가 익히기에는 다소 고난도이기 때문에 비전공자 혹은 어린이가 이를 배우고 구현하는데 프로그래밍 언어 학습과 개발 환경 설정 등의 과정이 지나치게 복잡할 수 있다[18]. 이러한 어려움으로 인해 많은 학습자가 머신러닝 학습을 포기하거나 학습 의지가 꺾이는 결과가 생기기도 한다.

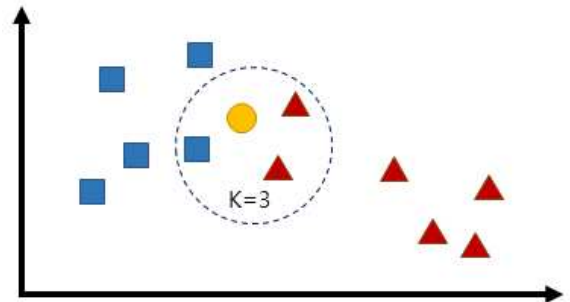
한편, 교육부(2020)에서는 인공지능 시대 교육정책의 방향과 핵심과제를 제시하면서, 유치원 때에는 게이미피케이션(Gamification) 기법을 활용하여 놀이를 통한 인공지능 교육을 실시하고, 초·중학생의 경우 수준에 맞는 인공지능 원리를 중심으로 교육하며, 고등학생의 경우 2021년 2학기부터 진로 선택 과목으로 'AI 기초'와 'AI 수학' 과목을 신설하여 인공지능 교육을 학교에 단계적으로 도입할 것을 밝혔다[12]. 이와 관련한 상세한 교육과정은 아직 발표되지 않아 머신러닝 관련 교육이 어느 단계에 어떻게 적용될지는 미지수이지만, 머신러닝은 인공지능 원리에 포함되므로 초·중학생부터 머신러닝 원리 학습이 구체적으로 실시될 것으로 예상할 수 있다.

이에 본 논문에서는 머신러닝의 원리 중 비교적 간단한 K-NN 원리를 활용하여 초등학생을 대상으로 머신러닝의 원리를 학습할 수 있는 교육 프로그램을 개발하고, 이를 시범적으로 적용한 결과를 바탕으로 효과성 분석을 실시하여 제안된 교육 프로그램의 현장 확대 적용의 긍정적 가능성을 타진해보고자 한다.

2. 관련 연구

2.1. K-NN 알고리즘

K-NN 알고리즘은 머신러닝 지도 학습 알고리즘의 한 종류로서 분류(classification)와 회귀(regression) 예측 문제에서도 사용되나 주로 분류 알고리즘으로 사용된다. K-NN은, K-Nearest Neighbors의 약자로 최근접 이웃 알고리즘으로 불리는데, 이는 새로운 데이터에 대해 기존 데이터 중 가장 가까운 K개 이웃 데이터의 정보를 통해 새로운 데이터의 정보를 예측하는 방법론을 의미한다[24]. Fig. 1은 K-NN 알고리즘의 기본 원리를 보여주고 있는데, 본 그래프에서 사각형은 정상 데이터, 삼각형은 비정상 데이터를 나타낸다고 할 때, 가운데의 원형은 주변 K개의 숫자에 따라 그 정보를 예측할 수 있다는 것이다. 만약 K가 3이라면 주변 데이터 3개의 정보에 따라 비정상 데이터로 예측할 수 있다[8].



(Fig. 1) K-NN Algorithm

K-NN에서 샘플과 샘플 사이의 거리를 구하는 계산식은 유클리디안 거리(Euclidean Distance) 방식을 활용한다. (1)은 AI 분야에서 여러 속성값에 의한 개체 사이의 유사도를 구할 때 사용하는 유클리디안 거리 계산식이다.

$$d(p,q) = d(q,p) = \sqrt{(q_1 - p_1)^2 + (q_2 - p_2)^2 + \dots + (q_n - p_n)^2} \quad (1)$$

$$= \sqrt{\sum_{i=1}^n (q_i - p_i)^2}$$

K-NN은 특히 유사도 기반의 단순한 자료의 매칭 방법으로 예측 결과를 도출하여 연산 속도가 빨라 결과 도출 시간이 빠르다는 장점이 있으며, 단기 교통상황 예측이나 개인화된 특성 예측 시간을 단축시킬 수 있어 인공지능 내비게이션에서 자주 사용되고 있다[7].

2.2. 머신러닝 교육 프로그램

머신러닝 교육에 있어 Amy J. Ko(2017)는 먼저 교육학적 내용 지식(Pedagogical Content Knowledge)의 관점에서 머신러닝을 학습해야 한다고 강조하였다[21]. 그는 ‘유용한 머신러닝 개념 표현’, ‘효과적인 머신러닝 개념 비유법, 예시 및 설명’, ‘머신러닝의 개념의 난점과 그 이유’, ‘머신러닝 학습 중의 실수’ 등을 학습하여, 학습자가 이를 자신의 언어로 표현할 수 있을 때 머신러닝을 보다 효과적으로 학습할 수 있다고 주장했다.

구글의 TensorFlow에서는 TensorFlow를 활용한 머신러닝의 기초 교육 프로그램을 제공하고 있다[26]. TensorFlow는 머신러닝 모델을 쉽게 개발하고 배포할 수 있도록 하는 플랫폼이다. 본 프로그램은 컴퓨터 공학 및 개발 관련 배경지식이 있는 학습자를 대상으로 하며, ‘머신러닝 이해’, ‘심화 학습’, ‘연습하기’, ‘TensorFlow 심화’의 4단계로 이루어져 있다.

국내 IT교육기관인 비트교육센터에서는 Python을 활용한 머신러닝으로 머신러닝 입문 기초과정을 제공하고 있는데, 주요 교육 주제로는 ‘머신러닝 개요와 기초 지식’, ‘Python의 패키지를 활용한 데이터 가공과 시각화’, ‘분류형 머신러닝의 기초 로지스틱 회귀의 원리’, ‘KNN, 랜덤 포레스트 등의 알고리즘 적용법’ 등 총 4차시의 교육과정으로 구성되어있다[22]. 한편, 이세훈, 박정준, 이명성(2020)의 연구에서는 머신러닝 학습을 위하여 Deep AI Yourself 실습 플랫폼을 개발하였다[17]. 본 프로그램은 스크래치와 TensorFlow를 접목하여 학습자가 블록 코딩을 활용해 인공지능 모델을 만들어볼 수 있도록 한다.

상기 언급한 교육 프로그램들은 기초과정이라고 하나 머신러닝 및 컴퓨터 프로그래밍 등의 기본 지식이 없는 학습자의 경우 난해할 가능성이 있다. 또한, 어린이 학습자를 위한 머신러닝 원리 교육 프로그램이나 언플러그드나 비대면 학습자료 등 다양한 형태의 교수학습자료에 관한 연구는 매우 미비하다. 인공지능 소양이 부족한 상태에서 단편적인 교육 실행은 사상누각(沙上樓閣)이 될 수 있다. 인공지능 전문가로 성장하거나 인공지능을 잘 활용할 수 있는 인재로 거듭나기 위해서는 머신러닝에 대한 기본 지식을 습득하는 것이 중요하다는 것을 알 수 있으며, 어린 학습자를 위한 머신러닝 교

육이 체계적이고 다양한 방식으로 실시되어야 할 것이다.

3. K-NN 알고리즘 이해를 기반한 머신러닝 교육 프로그램 설계

3.1. 교육 프로그램 개발 방향

본 교육 프로그램은 인공지능이 동작하기 위한 머신러닝 알고리즘 중 K-NN를 주제로 선정하여 개발되었다. 이에 본 논문은 머신러닝 K-NN 원리 교육 프로그램의 개발 방향을 다음과 같이 설정하였다.

첫째, K-NN의 기술적 이해나 문법 위주의 수업을 지양하고 원리 교육을 통한 인공지능 및 머신러닝 리터러시 함양을 중심으로 개발한다. K-NN은 초등학생이 기술적으로 이해하기는 난해하지만, 그 개념이 비교적 단순하여 어린이가 인공지능에 대한 접근성을 강화할 수 있는 소재이기 때문에 교육의 주제로 선정하였다.

둘째, 초등학교 고학년 실과 수업에서 활용할 수 있도록 구성한다. 초등학교 5~6학년의 2015 실과 개정 교육과정에서는 ‘소통’과 ‘혁신’ 개념에서 인공지능에 관한 소프트웨어와 하드웨어적 측면을 모두 다루고 있다[6]. 본 수업 진행 시 활용할 수 있는 자료로 구성한다.

마지막으로, 어린이 학습자임을 고려하여 놀이를 통한 배움이 가능하도록 구성한다. 학습자의 흥미도를 향상시키기 위하여 게임적 요소를 접목한 교육 방식은 학습자 스스로 교육에 능동적으로 참여할 수 있는 동기가 될 수 있을 것이다[10][11].

3.2. 교육 프로그램 개발 방법 및 절차

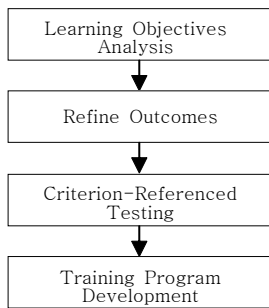
머신러닝 K-NN 원리를 활용한 교육 프로그램은 Robert F. Mager(1962)가 제안한 Criterion-Referenced Instruction(CRI)의 4단계 개발절차를 활용하여 본 프로그램에 맞게 세부 절차를 수정하여 사용하였다[16]. 첫 번째 단계에서는 무엇을 배울 것인지 교육의 주제와 학습 목표를 정의하고 구체화한다. 제안된 교육 프로그램은 머신러닝 K-NN 원리 학습을 바탕으로 궁극적으로 학습자의 머신러닝 원리 및 기술에 대한 관심과 교육적 호기심을 향상시키고 이를 통한 창의적 문제해결력을

증대시켜 미래형 창의 정보기술 인재 양성에 이바지하기 위한 목적으로 구성한다.

두 번째 단계에서는 학습을 통해 산출될 결과물을 구체화한다. 교육 프로그램은 학생들이 진행하면서 사용하는 학습지와 K-NN차트지, 자기 평가지가 산출물로 간주된다. 본 산출물은 이해도에 대한 평가를 목적으로 하지 않고 학습을 돕는 교육자료로서의 역할을 한다.

세 번째 단계에서는 명시된 결과의 목표를 달성하는데 필요한 지식이나 기술을 결정하여 프로그램을 구성하고 이를 평가할 방법을 정의한다. 첫 번째 단계에서 명시한 학습 목표에 따라 K-NN 원리 및 인공지능과 머신러닝에 대한 지식을 평가할 수 있는 퀴즈를 통해 학생들의 성취도를 평가한다.

마지막 단계는 첫 번째 단계에서 설정한 목표에 맞는 프로그램을 개발하는 것이다. 제안하는 프로그램은 초등학교 고학년을 대상으로 실과 및 창의적 체험학습 시간에 활용할 수 있는 총 10차시의 교육과정으로 구상하였다. Fig. 2는 프로그램의 개발절차를 시각화하여 보여준다.



(Fig. 2) Machine Learning Training Program Development Process

3.3. 교육 프로그램 설계

교육 프로그램은 학습자의 일상생활에서 기반한 문제에 대해 다룰 때 학습자의 수업 몰입도 향상을 기대할 수 있다[14][20]. 이에 본 논문에서 제안하는 교육 프로그램은 K-NN 알고리즘을 활용하여 인공지능 내비게이션이 길을 추천하는 원리를 파악하고 실제로 이를 구현해봄으로써 실생활에서 자주 접하는 문제에 대한 창의적 문제해결력을 기를 수 있다.

단계별 교육 프로그램의 내용은 이승철, 김태영

(2020)이 제안한 초등학생을 위한 인공지능 교육내용을 머신러닝 교육에 맞게 수정하여 활용하였으며, 자세한 내용은 Table 1에서 확인할 수 있다[19]. 먼저, 머신러닝의 개념에 대한 이해를 기본으로 머신러닝 개발에서의 윤리적 가이드라인을 이해하여 윤리적인 머신러닝 개발의 필요성을 인식한다. 이러한 기초지식을 학습한 후, 지도 학습 알고리즘의 여러 종류에 대한 이해도를 제고하여 K-NN 알고리즘의 개념, 장단점, 활용 분야 등을 학습한다. 이러한 배경지식은 K-NN을 활용하여 인공지능 내비게이션이 되어보는 실습 활동을 통해 머신러닝 적용 학습으로 마무리하게 된다.

<Table 1> Machine Learning Training Program based on Understanding K-NN Algorithm

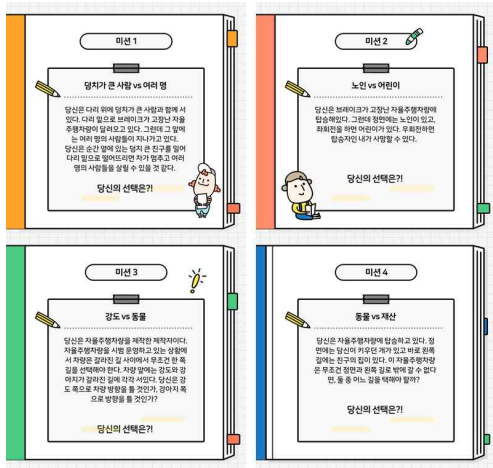
Class	Factor	Contents
1-3	Understanding ML principles	·ML Concepts ·ML Guiding Principles ·ML & Ethics
		·Supervised Learning Algorithms
4-7	Understanding K-NN Algorithm	·K-NN Algorithm Concepts ·Pros & Cons of K-NN Algorithm ·Fields where K-NN Algorithm is used
		·Problem Solving using ML - Principle of AI Navigation Recommends Directions
8-10	Application of ML	

한편, 제안된 교육 프로그램은 언플러그드, 놀이, 프로젝트, 문제기반학습, 토론, 실습 등의 다양한 교육 기법을 활용하여 수업에 적용하도록 한다. 초등학생 대상의 수업이기에 흥미 위주의 요소를 적절히 배치하여 학습자가 수업에 집중할 수 있도록 한다[2].

3.3.1. 머신러닝 원리 이해하기

‘머신러닝 원리 이해하기’ 단계는 제안된 교육 프로그램의 가장 첫 번째 단계로, ‘머신러닝 개념’, ‘머신러닝 개발자의 개발 원리’, ‘윤리적 머신러닝’을 주제로 총 3차시로 구성되어 있다. 머신러닝 개념 학습에서는 Tourestsky et al.(2019)이 제안한 K-12 인공지능 교육 가이드라인에서 착안해 초등학생이 알아야 할 머신러닝

개념에 관한 자료로 수업에 활용하였다[1]. 또한, 윤리적 머신러닝 수업에서는 인공지능이 가져올 대표적인 윤리 문제인 18)'트롤리 딜레마'를 비롯하여 자율주행차량이 마주할 여러 윤리적 문제에 관해 토론하여 머신러닝에서 어떠한 윤리적 접근이 필요한지를 학생 스스로 생각하면서 이를 표현해보도록 한다. Fig. 3은 2, 3차시에 활용되는 학습자료의 예시이다.

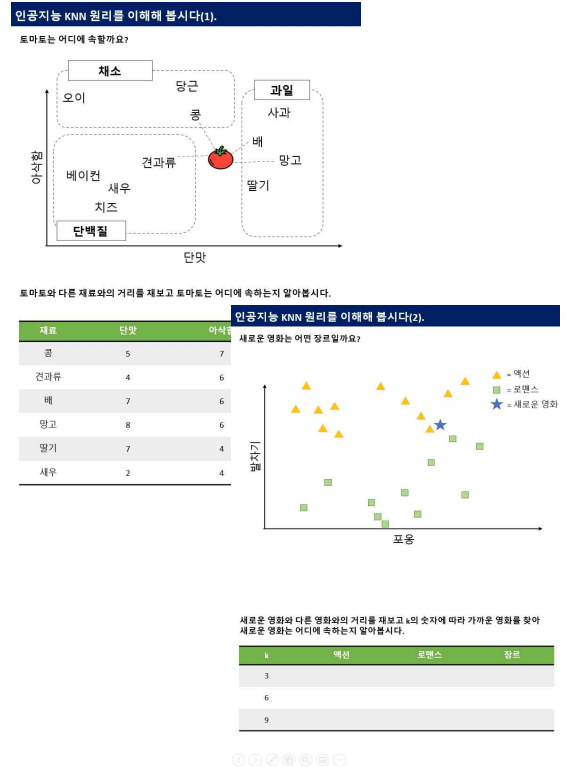


(Fig. 3) Educational Materials for Understanding ML Principles

3.3.2. K-NN 알고리즘 이해하기

‘K-NN 알고리즘 이해하기’ 단계는 제안된 교육 프로그램의 두 번째 단계로, ‘지도학습 알고리즘’, ‘K-NN 알고리즘의 개념’, ‘K-NN 알고리즘의 장단점’, ‘K-NN 알고리즘이 활용되는 분야’로 총 4차시의 수업으로 구성된다. 본 수업에서는 인공지능 내비게이션 실습에 들어가기 전 개념 다지기의 단계로 강의형 수업 기법과 놀이 기반 학습 기법을 적절히 활용하여 수업할 것을 제안한다. 특히, 학생들의 이해를 돕기 위해서 그래프를 통한 음식 분류, 영화 분류 등을 직접 체험해볼 수 있도록 한다. Fig. 4는 해당 단계에서 활용할 수 있는 교수-학습자료의 예시이다[23][25].

18) 트롤리 딜레마는 인공지능 자율주행차량의 판단 기준과 관련해 가장 대표적으로 언급되는 윤리 문제이다. Philippa Foot(1978)에 의해 처음으로 제기되었으며, 트롤리 방향에 따라 도덕적 선택의 기로에 서게 되는 여러 사례가 있다 [15].



(Fig. 4) Educational Materials for Understanding K-NN Algorithm

3.3.3. 머신러닝 적용하기

‘머신러닝 적용하기’ 단계는 제안하는 교육 프로그램의 가장 마지막 단계로 머신러닝을 활용하여 문제를 해결하는 학습을 진행한다. 해당 단계에서는 인공지능 내비게이션이 사람의 성향 및 특성에 따라 길을 추천해주는 원리를 실습한다. 해당 수업은 모둠 게임을 통해 K-NN 알고리즘의 원리를 이용한 머신러닝이 실생활에 적용된 사례를 학습할 수 있다. 게임에 앞서 준비물은 K-NN 차트지, 자기 평가지, 스마트폰, K-NN 학습 교구가 필요하다. Fig. 5는 게임에 필요한 자료의 예시이다. K-NN 차트지는 미리 교사가 데이터를 종합하여 작성한 자료를 바탕으로 하고, 자기 평가지에는 조급함, 검소함, 조심성, 하이패스 선호도, 과속 정도, 차 사고 현장 회피 정도, 위험 감수 정도의 기준에 따라 표시할 수 있도록 한다. 또한, 학습 교구의 경우 보드판 위에

도로를 그려놓았는데, 본 도로는 하이패스, 과속 CCTV, 차 사고 현장, 관광지, 넓고 안전한 길, 위험한 길 등의 특징이 있어 학생들이 게임을 하면서 도로의 다양한 상황에 맞게 길을 추천하는 원리를 학습할 수 있도록 구상한다[9].



(Fig. 5) Educational Materials for Application of ML

게임의 규칙은 다음과 같다.

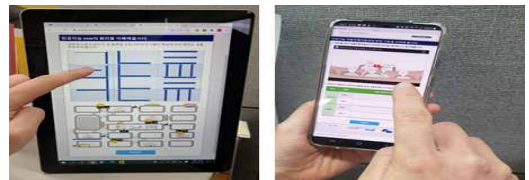
- 1) 모든 모듈원은 자기 평가표에 0-10점의 척도 안에서 본인의 성향을 표기한다. 예) 조급함 3, 하이패스 선호도 7 등
- 2) 술래를 정한다.
- 3) 술래는 지도를 보고 출발지점부터 완료지점까지 본인이 선호하는 길을 길 블록으로 표시한다.
- 4) 술래는 표시한 길을 스마트폰을 이용하여 사진을 찍어 놓고, 사진을 찍은 후 길 블록을 떼어 자신이 표시한 길을 다른 모듈원이 알 수 없게 한다.
- 5) 나머지 모듈원은 술래의 자기 평가표를 보고 K-NN 차트지에 술래의 값을 표시한다.
- 6) K-NN 알고리즘에 따라 표시한 값과 인접한 3개의 값을 확인하여 여러 종류의 길에 대한 술래의 선호도를 파악한다(K=3).
- 7) 파악한 선호도에 따라 길 블록으로 술래가 선택했을만한 길을 표시한다.
- 8) 스마트폰을 확인하여 술래가 선택한 길과 모듈

원이 예상한 길이 얼마나 차이가 있는지 확인한다.

9) 술래가 표시한 길과 정확하게 일치했다면 모듈원의 승리, 차이가 있다면 술래의 승리가 된다.

본 게임 규칙 중, 1, 5, 6번은 K-NN 차트지를 활용하여 데이터 샘플을 직접 표시함으로써 학생들이 주어진 K 값에 따라 K-NN 알고리즘이 데이터를 예측하는 방법을 몸소 파악할 수 있다. 또한, 나머지, 1, 2, 3, 4, 7, 8, 9번 규칙은 인공지능 네비게이션이 제시한 길과 운전자가 선호하는 길의 괴리감의 원인과 해결책에 대해 깊게 논의할 수 있는 단계로 구성되었다.

제안된 교육 프로그램의 특징은 대면 수업에서는 학습 교구를 통해 오감을 활용한 수업이 가능하고, COVID-19 감염병의 영향으로 비대면 수업에서도 활용할 수 있도록 온라인 학습지를 구상했다는 점이다. 온라인 학습지의 예시는 Fig. 6에서 확인할 수 있다.



(Fig. 6) Online ML Worksheets

4. K-NN 알고리즘 이해를 기반한 머신러닝 교육 프로그램 적용

4.1. 교육 프로그램의 타당도 분석

제안된 교육 프로그램의 타당도를 분석하기 위하여 초등학교 관리자와 인공지능·머신러닝 교육 전문가 15명을 대상으로 델파이 검사를 실시하였다. 타당도 분석을 위한 설문 문항으로는 류미영, 한선관(2019)이 제시한 인공지능 교육 프로그램의 타당성 문항을 활용하여 본 교육 프로그램에 맞게 수정하여 활용하였다[13]. 문항은 총 20문항이며, 6개의 요인으로 구분된다. 구분된 요인은 ‘머신러닝 교육 프로그램의 적절성’, ‘머신러닝 교육내용의 적절성’, ‘학생 학습 수준의 적절성’, ‘현장 수업 적용의 적합성’, ‘프로그램 내용과 교수-학습자료의 적절성’, ‘교육 프로그램 목표 달성 적합성’으로 구성하였으며, 결과는 Table 2와 같다.

<Table 2> Content Validity Ratio of Training Program (N=15)

Factor	CVR
1 Relevance of ML Training Program	.812
2 Relevance of ML Education Content	.793
3 Relevance of Students' Learning Levels	.695
4 Suitability of Field Class Application	.799
5 Appropriateness of Training Program Contents and Teaching-Learning Materials	.910
6 Suitability to Achieve Learning Objectives of Training Program	.612

텔레비전 조사에서 패널 수가 15명 이상인 경우, .60 이상의 CVR값이 나타나야 하는데, 조사 결과 모든 값이 .60 이상으로 나타나 교육 프로그램의 타당도가 적합한 수준으로 나타났음을 보여준다. 특히, 교육 프로그램 내용에 대한 교수-학습자료의 대응성이 높고 활용 가능성이 높다는 점에서 전문가들은 가장 높은 점수를 준 것으로 분석되었다. 한편, 교육 프로그램의 목표에 비해 K-NN 알고리즘 학습에 집중된 수업 내용이 아쉬웠다는 의견과 함께 프로그램의 목표 달성 적합도 요인의 결과값이 가장 낮게 집계되었다. 본 결과에 따라 2차 텔레비전 조사는 실시되지 않았으며, 향후 연구에서는 K-NN 뿐만 아니라 머신러닝의 다른 분야에 관한 교육 프로그램 내용도 다루는 편이 좋겠다는 의견을 수렴하여 더욱 다양한 내용의 프로그램 구성에 관한 연구가 시행되어야 함을 시사한다.

4.2. 교육 프로그램 적용의 효과 분석

본 논문에서 제안하는 교육 프로그램을 적용하고 프로그램의 효과를 분석하기 위하여 제주 J초등학교의 5학년 학생 25명을 대상으로 2주에 걸쳐 실시되었다. 시범 수업은 총 10차시 중 마지막 실습인 '머신러닝 적용하기' 3차시만 대면 수업으로 진행되었고, '머신러닝 원리 이해'와 'K-NN 알고리즘 이해' 수업은 비대면으로 진행되었다. Fig. 7은 진행된 수업 중 대면 수업의 모습을 보여준다.



(Fig. 7) Demonstration of the Training Program

시범 수업의 전과 후에 학생들의 머신러닝에 대한 이해도 차이를 평가할 수 있는 검사지로 학생들의 변화를 측정하여 t 검정으로 교육 프로그램에 대한 효과를 분석하였다. 본 검사지는 퀴즈의 형태로 출제되었는데, 본 평가 문항 요소는 교육 프로그램의 하위 요소와 상응된다. 크게 '머신러닝 원리'와 'K-NN 알고리즘'에 대한 이해, 그리고 '머신러닝의 적용'에 대한 이해도 평가가 가능하도록 구성한다. 전체 평가 문항은 총 8개로 Table 3과 같다. 검사 문항에 대하여 학생들은 Likert 5점 척도로 응답하였다.

<Table 3> Composition of Evaluation Questions

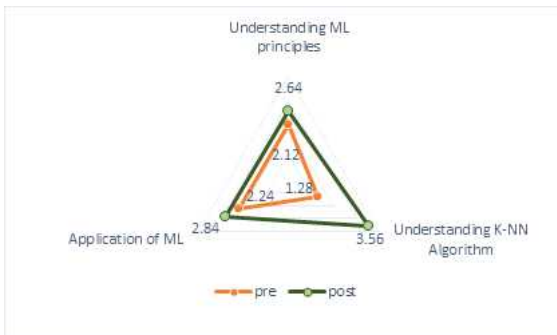
Evaluation Factor	Sub-Questions
Understanding ML principles	1. I understand ML concepts.
	2. I can identify ML guiding principles.
	3. I know what ethical ML is.
Understanding K-NN Algorithm	4. I can classify supervised learning algorithms.
	5. I understand K-NN algorithm concepts.
	6. I can represent pros and cons of the K-NN Algorithm.
	7. I can identify fields where K-NN algorithm is used.
Application of ML	8. I can solve the problems using ML.

상기 평가 문항으로 학생들의 사전·사후 점수 차에 의한 대응표본 t 검정의 결과는 Table 4와 같으며, 이를 사전 사후 점수를 중심으로 도식화한 결과는 Fig. 8과 같다. 3가지 요인 중 가장 통계적으로 유의미한 결과를 나타낸 요인은 'K-NN 알고리즘'에 대한 이해도 제고로 나타났다($p<.001$). 학생들은 수업 이전에는 거의 K-NN에 관해 들어본 적이 없는 친구들로 구성되었기 때문에 이와 같은 결과가 나타났을 것으로 파악된다. 또한, '머신러닝 원리'에 대한 이해와 '머신러닝의 적용'에 대한 이해도에서도 통계적으로 유의미한 차이를 나타내었다($p<.01$). 학생들은 머신러닝이라는 용어는 들어본 적이 있으나 동작의 원리나 실생활에 어떻게 활용되는지에 관해서는 수업 전에 인식하지 못했던 경우가 많아 머신러닝에 대한 전반적 인식 제고에 도움이 되었다고 볼 수 있다. 이러한 결과는 본 프로그램이 머신러닝에 관한 이해도 향상에 교육적 효과가 있음을 보여준다.

<Table 4> Understanding Machine Learning score differences in the pre and post test

Measure		M	SD	t
Understanding ML principles	pre	2.12	.666	-3.641**
	post	2.64	.638	
Understanding K-NN Algorithm	pre	1.28	.458	-11.225**
	post	3.56	.676	
Application of ML	pre	2.24	.663	-3.133**
	post	2.84	.800	

*** $p<.001$, ** $p<.01$, * $p<.05$



(Fig. 8) Comparison of Pre-Post Test Result of Understanding Machine Learning

5. 결론

본 논문은 K-NN 알고리즘 원리를 중심으로 머신러닝에 대한 이해도를 제고할 수 있는 교육 프로그램을 개발한 결과를 보여준다. 교육 프로그램은 '머신러닝 원리 이해하기', 'K-NN 알고리즘 이해하기', '머신러닝 적용하기'의 3단계로 총 10차시에 걸쳐 진행된다. 또한, 교육 프로그램을 효과적으로 적용할 수 있도록 다양한 창의적 기법을 활용하고 놀이 활동을 개발하여 낯설게 느껴질 수 있는 머신러닝 개념을 학생들로 하여금 흥미있게 학습할 수 있도록 하였다. 개발된 교육 프로그램은 전문가들에 의해 그 내용적 타당도가 검증되었으며, 초등학교 시범 적용을 통하여 교육 프로그램에 참가한 전반적인 학생들의 머신러닝에 관한 이해가 향상되었음을 알 수 있었다. 다만, 교육적 효과는 통계적으로 유의하였으나, 머신러닝에 대한 지식이 거의 없었던 학생들을 대상 시범 수업을 진행했기 때문에 'K-NN 알고리즘' 요인을 제외하고 다른 요인의 이해도 평가에서는 다소 저조한 점수가 집계되었다. 이는 K-NN 알고리즘 학습에 집중하여 학생들의 머신러닝의 다른 분야에 대해 깊게 이해하지 못했을 수 있다고 판단된다. 이는 전문가 타당도 검사에서도 관련 의견이 제기된 만큼, 향후 연구에서는 머신러닝의 다양한 분야를 다루는 교육 프로그램 개발이 진행되어야 할 필요가 있음을 시사한다.

본 논문에서 제안한 교육 프로그램은 다양한 인공지능 기기에서 활용되는 K-NN 알고리즘을 주제로 구상한 머신러닝 교육으로 학생들의 흥미를 불러일으킬 수 있는 교육 프로그램이다. 이는 상대적으로 미비한 초등학생 대상 머신러닝 원리 학습 교육 프로그램을 개발하였다는데 본 논문의 의의를 찾을 수 있으며, 향후 미래형 창의 정보 인재 양성과 인공지능 및 머신러닝 교육 프로그램 개발에 기여할 수 있기를 기대한다.

참고문헌

- [1] David S. Touretzky, Christina Gardner-McCune, Fred Martin and Deborah Seehorn (2019). K-12 Guidelines for Artificial Intelligence: *Proc. of the ISTE Conference*.
- [2] DongHyeok Lee, NamJe Park (2016). Teaching Book and Tools of Elementary Network Security Learning using Gamification Mechanism. *Journal of Korea Institute of Information Security & Cryptology*, 26(30), 787-797.
- [3] Donghyeok Lee, Namje Park (2017). Geocasting-based synchronization of Almanac on the maritime cloud for distributed smart surveillance. *The Journal of Supercomputing*, 73(3), 1103-1118.
- [4] Donghyeok Lee, Namje Park (2018). Electronic Identity Information Hiding Methods Using a Secret Sharing Scheme in Multimedia-centric Internet of Things Environment. *Personal and Ubiquitous Computing*, 22(1), 3-10.
- [5] HyungJoo Kim, ShinHyoung Park and Kitae Jang (2016). Short-term Traffic States Prediction Using k-Nearest Neighbor Algorithm : Focused on Urban Expressway in Seoul. *Journal of Korean Society Transportation*, 34(2), 158-167.
- [6] HyunJin Jang(2020). The Domestic Research Trends of Artificial Intelligence Related in Elementary Practical Arts Education. *Korean Journal of Elementary Education*, 31, 33-48.
- [7] JangYeol Lee, ChoonSung Nam and DongRyeol Shin (2017). *Machine Learning Bigdata Education Platform using Apache Spark*. Korea Computer Congress 2017 Program Book, 1531-1533.
- [8] Javed Asharf, Nour Moustafa, Hasnat Khurshid, Essam Debie, Waqas Haider and Abdul Wahab (2020). A Review of Intrusion Detection Systems Using Machine and Deep Learning in Internet of Things: Challenges, *Solutions and Future Directions*. *Electronics*, 9(7), 20-21.
- [9] JinSu Kim, NamJe Park (2019). Lightweight Knowledge-based Authentication Model for Intelligent Closed Circuit Television in Mobile Personal Computing. *Personal and Ubiquitous Computing*, 1-9.
- [10] JinSu Kim, NamJe Park (2019). Development of a board game-based gamification learning model for training on the principles of artificial intelligence learning in elementary courses. *Journal of The Korean Association of Information Education*, 23(3), 229-235.
- [11] JinSu Kim, Namje Park (2020). A Face Image Virtualization Mechanism for Privacy Intrusion Prevention in Healthcare Video Surveillance Systems. *Symmetry*, 12(6), 891
- [12] Ministry of Education(2020). Policy Directions and Key Tasks of the Artificial Intelligence Era.
- [13] MiYoung Ryu, SeonKwan Han (2019). AI Education Programs for Deep-Learning Concepts. *Journal of The Korean Association of Information Education*, 23(6), 583-890.
- [14] NamJe Park, YoungHoon Sung, YoungSik Jeong, Soo-Bum Shin and Chul Kim (2018). The Analysis of the Appropriateness of Information Education Curriculum Standard Model for Elementary School in Korea. *International Conference on Computer and Information Science*. Springer. 1-15.
- [15] Philippa Foot(1978). *The Problem of Abortion and the Doctrine of the Double Effect in Virtues and Vices*. Berkeley: University of California Press.
- [16] Robert F. Mager(1962). *Preparing Objectives for Programmed Instruction*. Belmont: Fearon Publishers.
- [17] SeHoon Lee, JeongJun Bak and MyeongSung Lee (2020). Educational Programming Language based Deep AI Yourself Hands-on Platform for Machine Learning. *Proceedings of the Korean Society of Computer Information Conference*, 28(2), 243-244.
- [18] SeHoon Lee, JiHyun Jeong, JinHyeong Lee and CheonWoo Jo(2020). D.I.Y : Block-based

Programming Platform for Machine Learning Education. *Proceedings of the Korean Society of Computer Information Conference*, 28(2), 245-246.

- [19] SeungChul Lee, TaeYoung Kim(2020). Proposal of Contents and Method of Artificial Intelligence Education for Elementary School Students. *Proceedings of The Korean Association of Computer Education*, 24(1), 177-180.
- [20] SungJin Hwang, JeongWon Choi and YoungJun Lee(2014). Development of Education Plan of App Inventor for Improving Informatics Gifted Elementary Students' Learning Flow. *Proceedings of The Korean Association of Computer Education*, 18(2), 158-190.
- [21] Amy J. Ko(2017). We need to learn how to teach Machine Learning. <https://medium.com/bits-and-behavior/we-need-to-learn-how-to-teach-machine-learning-acc78bac3ff8>
- [22] Bit Academy . http://www.bitacademy.com/Course/ShortCourse/DetailPage/Course_DSCIENCE_PracticalMachineLearningPython.asp.
- [23] Duy Anh Nguyen(2019). Classifying with K-Nearest Neighbors. https://medium.com/@duyanhnguyen_38925/classifying-with-k-nearest-neighbors-17e202da449
- [24] Onel Harrison(2018). Machine Learning Basics with the K-Nearest Neighbors Algorithm. <https://towardsdatascience.com/machine-learning-basics-with-the-k-nearest-neighbors-algorithm-6a6e71d01761>
- [25] SungHyun Ahn(2019). Mining - K Nearest Neighbors . <https://shacoding.com/2019/12/07/mining-k-nearest-neighbor/>
- [26] TensorFlow. <https://www.tensorflow.org/resources/learn-ml/basics-of-machine-learning?hl=ko>.

저자소개



최 은 선

2014 가천대학교 경영학과(학사)
 2016 북경어언대학교 국제중국어교육(석사)
 2020~현재 제주대학교 일반대학원 컴퓨터교육전공(박사과정)
 2020~현재 제주대학교 과학기술사회연구센터, 사이버보안인재교육원, 창의교육거점센터 선임연구원
 관심분야 : 융복합교육, 컴퓨터교육, 인공지능교육, 창의교육 등
 e-mail : choi910624@jejunu.ac.kr



박 남 제

2008 성균관대학교 컴퓨터공학과(공학박사)
 2003~2008 한국전자통신연구원 정보보호연구단 선임연구원
 2009 University of California at LA (UCLA) Post-doc
 2010 Arizona State University (ASU) Research Scientist
 2010~현재 제주대학교 교육대학 초등컴퓨터교육전공 교수, 과학기술사회(STS)연구 부센터장, 사이버보안인재교육원장, 창의교육거점센터장
 관심분야 : 컴퓨터교육, STEAM, 정보보호, 암호이론 등
 e-mail : namjepark@jejunu.ac.kr