Instructional Technology and Lifelong Learning Vol. 5, Issue 2, 23-45 (2024)

https://dergipark.org.tr/en/pub/itall

ITALL

ISSN: 2717-8307

Research Article

Can Machine Learning Be Taught to Pre-service Teachers in the STEM Fields?

Esma Nur Özen [1], Elif Polat *[2], Yavuz Samur [3]

* *Corresponding Author, elif.polat@iuc.edu.tr*
[1]*Istanbul University-Cerrahpasa, Türkiye*
[2]*Istanbul University-Cerrahpasa, Türkiye*
[3]*Bahçeşehir University, Türkiye*

Abstract

Machine Learning (ML) trainings provide students with 21st century skills and enable students to find solutions to their own problems. The purpose of this study is to design, implement, and evaluate ML training for pre-service teachers in the STEM field in order to contribute to the future workforce in the field of computer science. The participants of the study were 74 pre-service teachers who are studying in the departments of Computer Education and Instructional Technology (CEIT), science education, and mathematics education (STEM fields) at a state university in Istanbul. Convenience sampling method was used in the study. In the research, a single-group pre-test-post-test weak quasi-experimental design was used by using the quantitative method in order to evaluate the training by giving ML training. The training was implemented on the online platform for 3 hours for 8 weeks. "Pretest - Posttest Achievement Test," "Online Student Engagement Scale," "Moodle Activity Data," "Demographic Form," and "Attendance Forms" were used to collect data. There is a significant difference between the pre-test and post-test averages in favor of the post-test. There is a significant difference between the pretest and posttest scores according to the departments. It has been concluded that the provided training is effective in the success of pre-service teachers. It can be suggested to offer training to different branches and to select participants from elementary and middle school students.

Keywords: Machine learning, Machine learning instruction, STEM, artificial intelligence, preservice teachers.

Citation: Özen, E. N., Polat, E., & Samur, Y. (2024) "Can machine learning be taught to pre-service teachers in the STEM fields?". *Instructional Technology and Lifelong Learning, 5(2),* 23 - 45. *https://doi.org/10.52911/itall.1458322*

STEM Alanındaki Öğretmen Adaylarına Makine Öğrenmesi Öğretilebilir mi?

Özet

Makine öğrenmesi eğitimleri, öğrencilere 21. yüzyıl becerileri kazandırır ve kendi problemlerine çözüm bulmalarını sağlar. Bu çalışmanın amacı, bilgisayar bilimi alanında gelecekteki iş gücünün oluşturulmasına katkıda bulunmak amacıyla STEM alanındaki öğretmen adaylarına yönelik makine öğrenmesi öğretimini planlamak, uygulamak ve değerlendirmektir. Çalışmanın katılımcıları, İstanbul'da bir devlet üniversitesinde 2020-2021 akademik yılında bilgisayar ve öğretim teknolojileri eğitimi, fen eğitimi ve matematik eğitimi (STEM alanları) bölümlerinde öğrenim gören 74 öğretmen adayıdır. Çalışmada elverişli örnekleme yöntemi kullanılmıştır. Araştırmada, makine öğrenmesi eğitimi verilerek eğitimin değerlendirilmesi amacıyla nicel yöntem kullanılarak tek gruplu ön-test-son-test zayıf yarı deneysel tasarımı kullanılmıştır. Eğitim, 8 hafta boyunca çevrimiçi platformda haftada 3 saat olacak şekilde uygulanmıştır. Veri toplama araçları olarak "Ön Test - Son Test Başarı Testi", " Çevrimiçi Öğrenci Bağlılık Ölçeği", "Moodle Etkinlik Verileri", "Demografik Form" ve "Katılım Formları" kullanılmıştır. Ön-test ve son-test ortalamaları arasında son-test lehine anlamlı bir fark vardır. Bölümlere göre ön test ve son test puanları arasında anlamlı bir fark bulunmaktadır. Verilen eğitimin öğretmen adaylarının başarısında etkili olduğu sonucuna ulaşılmıştır. Farklı branşlara eğitim vermek ve katılımcıların ilkokul ve ortaokul öğrencilerinden seçilmesi önerilebilir.

Anahtar Kelimeler: Makine öğrenmesi, makine öğrenmesi eğitimi, STEM, yapay zeka, öğretmen adayları.

Date of Submission	25.03.2024
Date of Acceptance	04.07.2024
Date of Publication	31.12.2024
Peer-Review	Double anonymized - Two External
Ethical Statement	It is declared that scientific and ethical principles have been followed while carrying out and writing this study and that all the sources used have been properly cited.
Acknowledgements	This article is extracted from my master thesis/doctorate dissertation entitled “Planning, implementing and evaluation of ML training for preservice teachers in STEM field”, supervised by Yavuz Samur & Elif Polat
Author(s) Contribution	**Author1:** Conceptualization, Methodology, Writing- Original draft preparation. **Author2:** Conceptualization, Supervision, Reviewing and Editing **Author3:** Supervision, Reviewing and Editing.
Plagiarism Checks	Yes - Turnitin
Conflicts of Interest	The authors have no conflict of interest to declare.
Complaints	itall.journal@gmail.com
Grant Support	The authors acknowledge that they received no external funding in support of this research.
Copyright & License	Authors publishing with the journal retain the copyright to their work licensed under the **CC BY 4.0**.

1. Introduction

With the tools that technology provides, we can complete many of our daily tasks in a faster, easier, and more practical manner. When used in education, it not only facilitates tasks but also enhances learning. Teachers' technological knowledge is critical at this point. The use of artificial intelligence (AI) in education has increasingly become more widespread. AI is referred to as the programming of human intelligence functionalities to solve problems on a computer (Nabiyev, 2012). In the field of education, AI is used in profile creation and estimation, assessment and evaluation, adaptive systems, personalized learning, and intelligent teaching systems (Senocak, 2020). Machine learning (ML), a subset of AI, is also finding applications in education. ML is a method of determining a solution by predicting the information taught by the computer in response to a future situation (Çevik & Kayakus, 2020). Many countries have begun to employ ML practices to aid in student development. In education, AI and ML practices are used to provide feedback for individualized instructional plans (Kayahan, 2018). Using AI and ML, it is possible to adapt to a changing world because special educational content can be prepared for each individual student (Demirkaya & Sarpel, 2018). The use of ML is crucial at this point. ML is also used to attract student attention and reduce teacher workload (Nafea, 2018). ML is used to recognize faces/voices, determine success, and separate students based on certain criteria. Another application is to estimate why students drop out of school and identify risks at school (Mduma et al., 2019). Murphy (2019) notes that systems exist which enable educators to proactively intervene by alerting them in advance if a student is expected to be absent on a future date. To summarize, ML platforms are used in education to teach students ML, classify data, and make predictions in general.

Knowledge of mathematics and computers is of great importance in learning ML (Reyes et al., 2020). In the realm of education, ML finds application across diverse areas such as assessment and evaluation, predicting achievement, creating course content, identifying learning styles, and developing intelligent course systems. However, research is insufficient to keep up with the rapidly changing and developing field of education. According to several research studies, AI and ML are causing the birth of new professions. Many countries have shifted their educational policies to avoid unemployment, and Albania sees AI as a business opportunity in this regard (Tataj & Kola, 2021). As a requirement of computer science, learning AI and ML emerges as a necessity (Chung & Shamir, 2020).

According to Sakulkueakulsuk et al. (2018), mastering and applying AI along with its subdivisions is both labor-intensive and challenging. Teaching ML as part of an AI course in undergraduate and graduate schools helps shape career goals. It has been stated that providing AI and ML education programs early enables students to develop career plans at a young age. Based on the studies on the STEM approach, STEM literacy training programs should be planned. As mentioned by Bybee (2010), preservice teachers and in-service teachers should be trained to equip people with 21st-century skills. The high level of knowledge that teachers have in the STEM approach and the development of a perspective towards the STEM approach enhance the efficiency of STEM education (Wang et al., 2011).

Kim and Kwon (2024) conducted a systematic review of 36 articles evaluating AI education at the K-12 level from 2013 to 2022. They found that introducing students to AI assessments in formal learning environments from an early age is crucial. Martin et al. (2024) organized five projects as exhibitions to help 125 elementary and middle school students aged 7-14 understand AI and ML. The findings showed that students were able to process data from cameras involved in ML and respond to the system's confidence intervals. Not only do programming skills improve as a result of training, but so do creativity and cognitive thinking skills.

There are benefits to programming education, but to translate these benefits into practice, tools appropriate for the target audience should be chosen. Therefore, a platform with AI and ML that is suitable for the target audience should be selected. For a target audience with basic programming knowledge, block-based ML platforms are preferred, whereas text-based ML platforms are preferred for a target audience with intermediate and higher knowledge. Quiroz and Gutierrez (2024) developed activities using a Scratch extension to study middle school students' experiences. They found that students' interest in AI increased and the coding foundation offered was more beneficial compared to other technological tools. Priya et al. (2024) conducted an experimental controlled study with 41 high school students by developing a game with ML. The students in the experimental group performed better in tests than those in the control group, showing that the game helped introduce ML concepts.

According to a study conducted by Tektas et al. (2010), AI courses should be included as compulsory courses at the undergraduate and graduate levels. Young people educated in AI and ML contribute to a significant decrease in import purchases of countries. In this context, teaching ML is a requirement for the future workforce. It is worth noting that text-based Python and the block-based ML for Kids platform are being used to teach ML. It was found that the ML for Kids platform is not yet widely used in education (Zhu, 2019). In a study conducted by Park et al. (2020), the ML platform called ML for Kids was used to improve the instructional model for software education based on ML.

With the growing importance of ML, how to incorporate it into computer science education is viewed as a problem that field experts and practitioners should work on collaboratively in an interdisciplinary manner (Zhu, 2019). In education, ML is used for teaching, classification, and estimation purposes. ML training is becoming more popular as the demand for up-to-date educational content, technical support, and educational content that is relevant to everyday life grows.

Tseng et al. (2024) organized a two-week AI and ML Summer Camp for teenagers aged 13-18 to create ML-supported personal mobile applications in teams. This study highlighted the importance of collaboration, model testing interfaces, and student-centered projects in actively engaging students in exploring the role of data in ML systems. In another study, Reyes et al. (2020) trained high school students in ML by adding ML training to curricula through gamification, aiming to teach students the fundamentals of ML. The students expressed their pleasure in experiencing ML through Scratch. Moreover, despite the fact that the majority of the participants had prior knowledge of ML, it was concluded that learning ML was simple because complex concepts were not part of the training.

Peters (2019) aimed to teach the principles of programming to high school students through ML. Working with robots that were programmed using ML was found to be motivating, and the workshop, as a group activity, was a lot of fun.

Students can gain 21st-century skills by participating in training programs that introduce ML and explain how to use it. As a result, students can be empowered to solve their own problems. For this reason, training should be given to teachers in the STEM field as early as possible. Moreover, it has been found that the literature lacks studies that contribute to the use of ML in education. The literature shows that studies on ML training are limited. Therefore, ML training programs should be offered to contribute to the literature and to the development of preservice teachers in science education, mathematics education, and CEIT departments, which are some of the STEM fields in faculties of education. Preservice teachers may be given new opportunities if the training programs are related to the technology that is considered necessary today. Experimental studies should be carried out on this subject, and how to use it concretely in education should be examined (Demirkaya & Sarpel, 2018).

For this reason, in the present study, preservice teachers in the STEM field were given training in ML, and the training was evaluated. The goal of this study was to plan, implement, and evaluate the ML training for preservice teachers in the STEM field in order for them to contribute to the development of the computer science field's future workforce. In this context, the research questions of this study are as follows:

- Is the ML training provided to pre-service teachers in STEM fields effective in teaching ML?
- Is there a significant difference between pre-test and post-test achievement test scores across departments?
- Is there a significant difference between pre-test and post-test achievement test scores based on practical and theoretical questions in the achievement test?
- To what extent does engagement in the ML training given to pre-service teachers in the STEM field predict achievement?

2. Method

2.1. Research Model

In this study, the ML training was given to preservice teachers in a faculty of education's CEIT, science education, and mathematics education departments, and the training was evaluated using quantitative methods. Quantitative methods are research methods that rely on objective measurement and analysis (Buyukozturk et al., 2008).

This study was prepared as a single-group pretest-posttest quasi-experimental design. Single-group pretest-posttest models are impartial implementations of pre- and post-practice measurements (Karasar, 2017). Single-group pretest-posttest models are used to conduct pre- and post-practice analyses. This study made use of an achievement test (as a pretest and a posttest), Moodle activity data, the Demographic Data Form, the Course Attendance Form, and the Online Student Engagement Scale.

2.2. Sample

The study was conducted on a completely voluntary basis. Special announcements about the training were made to preservice teachers studying in the departments of science education, mathematics education, and CEIT.

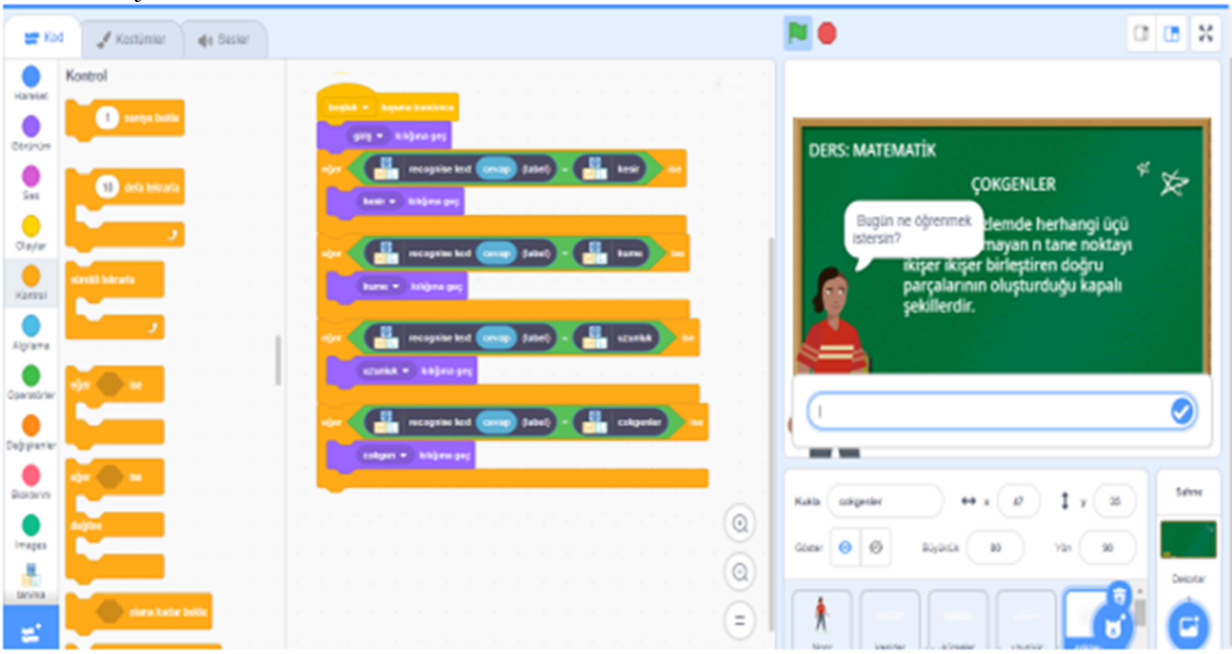
The convenience sampling method was used in the study. This sampling method is the most efficient and straightforward way to select a sample in order to avoid issues with funding, time, and workload (Buyukozturk et al., 2008). The sample consisted of 74 preservice teachers who were studying in the departments of CEIT, science education, and mathematics education at a state university in Istanbul.

2.3. Training Material Preparation Process

The researcher reviewed the literature for examples in the field of practice before preparing the training content. Due to the limited Scratch activities regarding ML for Kids, the training content was prepared by the researcher. Some activities were adapted from the existing activities on the ML for Kids platform. A total of 17 activities were developed for the training, covering the subjects of text, numbers, audio, and images. Additionally, a presentation was prepared to present general information after a literature review was carried out.Two educational technology specialists and two computer engineers working in the field of AI and ML provided expert opinions during the content preparation process. In addition, preservice teachers were interviewed to ensure the content was appropriate for the target audience. Great care was taken to create content that was relevant and suitable for the preservice teachers.

Figure 1

ML Activity



2.3.1. ML Training Process

Preservice teachers in the departments of CEIT, science education, and mathematics education received the training over an eight-week period. The ML training included the use of the ML for Kids and Scratch 3.0 platforms. ML for Kids is a free ML platform designed for children (Lane, 2018). Text, numbers, sound, and images are used to create projects on this platform. Block- and text-based ML projects can be created on the Scratch, Python, and App Inventor platforms (Zhu, 2019).

Since the participants would be learning about ML platforms for the first time during the training, the preparation of ML projects using the Scratch platform, a block-based coding platform, was approved by the experts. Before the training began, the participants were given general information about the training and asked to participate voluntarily. The importance of attending the training regularly was emphasized, and then the Demographic Data Form and the achievement test as a pretest were administered through an online platform. Due to the Covid-19 pandemic, the training was delivered as a live course via an online platform. The first lesson was planned as a general introduction.

Table 1

Machine Learning Training Plan

Week	Seminar	Topic(s)	Activity
1	-	General Information and Presentation	What is Artificial Intelligence? The Future of Artificial Intelligence History of Artificial Intelligence Types of Artificial Intelligence What is Machine Learning? Methods of Machine Learning Usage Areas Introduction to Machine Learning Advantages of Machine Learning Deep Learning (**180 min**)
2	Fundamentals and Current Topics of Machine Learning (**60 min**)	Texts	Learning Emotions Animal Kingdom I Want to Learn Mathematics Tourist Information Chatbot (**120 min**)
3	Use of Adaptive Systems in Education (**45 min**)	Numbers	Eat the Bug Pacman Game (**135 min**)
4	Artificial Intelligence and Fuzzy Technology (**45 min**) Use of Smart Interfaces in Education: Examples from the Artificial Intelligence Center (**30 min**)	Numbers	XOX Game Book Recommendation (**105 min**)
5	Artificial Intelligence and Robotics (**60 min**)	Sounds	Recognize My Voice Open Sesame Listen and Explain Our Target is the Bottle (**120 min**)
6	Modeling Metacognitive Activities in Informal Learning Environments: Recommendations from Theory to Practice (ML) (**45 min**)	Images	Recognize My Face Book Analysis My Best Friend Tom Rock-Paper-Scissors Game Making (**135 min**)
7	Use of ML Techniques in Education (**45 min**)	-	Project Presentations (**135 min**)
8	Artificial Intelligence in Education (**45 min**) Applications of Artificial Intelligence Algorithms (**30 min**)	-	Project Presentations (**105 min**)

2.4. Data Collection

The preservice teachers were given the achievement test online as a pretest before the ML training and as a posttest after the training. The Demographic Data Form was administered prior to the training, and the Online Student Engagement Scale was administered afterwards, both of which were done online. The participants' engagement was closely monitored to ensure that they did not drop out during the training process. The purpose and content of the training, as well as the voluntary nature of participation, were explained to the participants prior to the training.

2.5. Data Collection Instruments

In this study, the achievement test (as a pretest and a posttest), Moodle activity data, the Demographic Data Form, the Course Attendance Form, and the Online Student Engagement Scale were used as data collection instruments.

2.5.1. Achievement Test (Pretest and Posttest)

The achievement test, which included the subjects found in the ML curriculum for preservice teachers in the STEM field, was developed by the researcher. An indicator chart with eight main objectives was created. To assess these objectives, 38 multiple-choice questions were formulated. Necessary changes were made, and a pilot study was carried out with ten preservice teachers. Subsequently, item difficulty (p) and item discrimination (r) were analyzed.

Table 2

Item Analysis Results for The Achievement Test (Pretest and Posttest)

Item Difficulty	Item Discrimination
0.72	<0.20 and >0.40

According to Table 2, there were 30 questions with high item discrimination ($r > 0.40$). There were 6 questions with item discrimination between 0.25 and 0.39, and 2 questions with item discrimination of $r < 0.20$. Item difficulty was 0.75 for 32 of the questions. As a result, the average item difficulty (p) was found to be 0.72. The developed achievement test was determined to be appropriate based on the analysis.

2.5.2. Online Student Engagement Scale

The original form of the Online Student Engagement Scale was developed by Dixson (2015) and adapted into Turkish by Polat et al. (2022). The 19-item scale had 5-point Likert-type options. These were: (1) Does not define me at all, (2) Does not define me, (3) Undecided, (4) Defines me, and (5) Defines me completely. The scale consisted of 4 factors: skills (6 items), emotion (5 items), engagement (6 items), and performance (2 items).

Table 3

Reliability Results for The Online Student Engagement Scale

Cronbach alpha internal consistency coefficient	χ^2	SD
0.95	273.844	142

The Cronbach alpha internal consistency coefficient ranged from α = 0.77 to α = 0.87. The $\chi 2$ value was 273.844, and the standard deviation (SD) was found to be 142. Accordingly, the $\chi 2$ / SD (273.844 / 142) ratio was found to be 1.928. Cronbach alpha was determined to be 0.95. In the current study, the Cronbach alpha internal consistency coefficient was found to be 0.95 as well. As a result, it was determined that using the scale was appropriate.

2.5.3. Moodle Activity Data

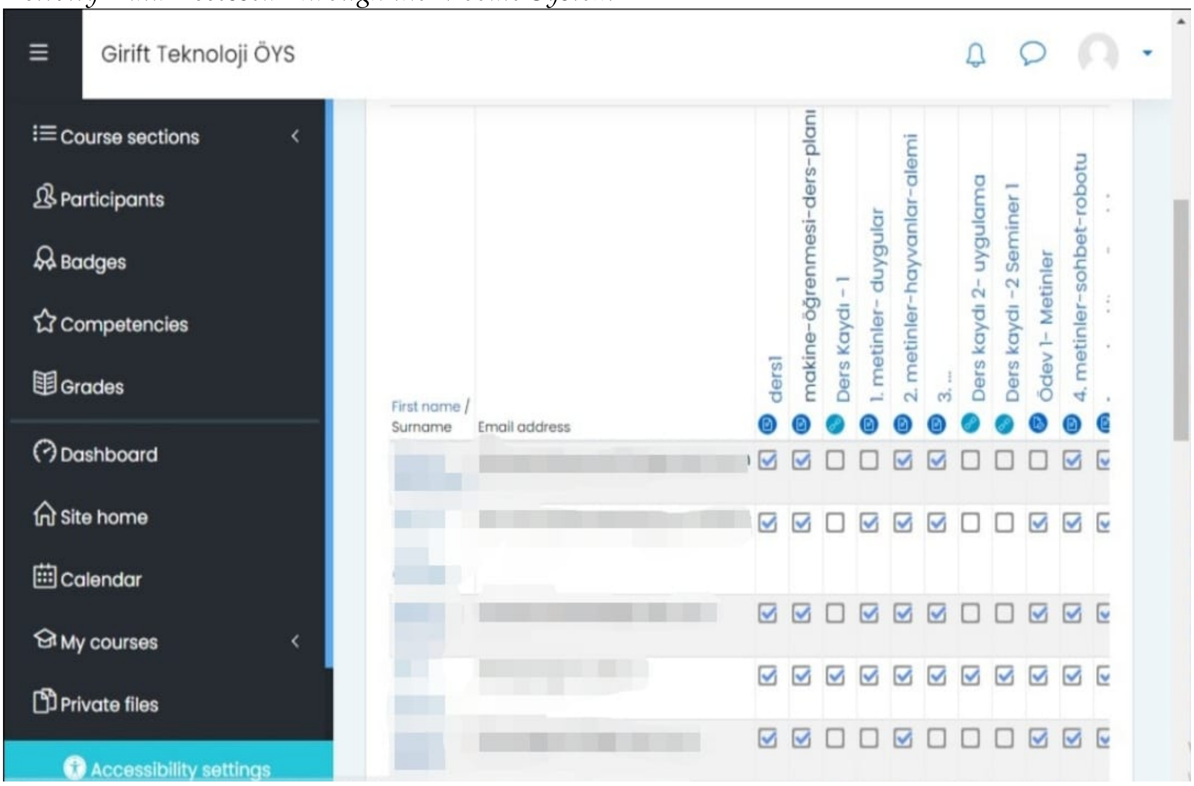
Data were gathered to determine the level of engagement in the training by the participants. Moodle activity data were collected on how participants reviewed course content and course recordings after they were shared in the Moodle system. Figure 2 shows the content shared within the Moodle system. The system automatically marked the status of the participants who examined the content as "completed," as shown in Figure 3.

Figure 2

Content Shared Within the Moodle System

Figure 3

Activity Data Accessed Through the Moodle System

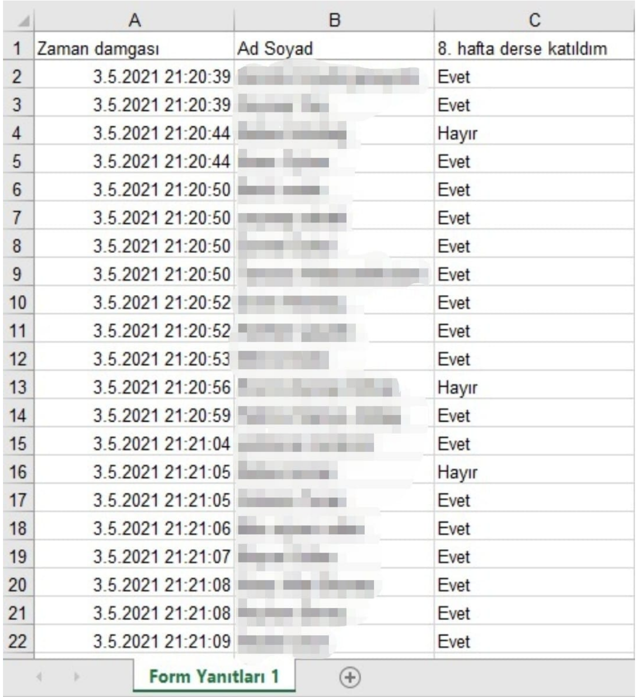


2.5.4. *Course Attendance Form*

The Course Attendance Form was used to track who was present during the training. It was administered through the online platform during each training session. The Course Attendance Form taken via the online platform is shown in Figure 4.

Figure 4

The Course Attendance Form Taken from The Online Platform



	A	B	C
1	Zaman damgası	Ad Soyad	8. hafta derse katıldım
2	3.5.2021 21:20:39	[illegible]	Evet
3	3.5.2021 21:20:39	[illegible]	Evet
4	3.5.2021 21:20:44	[illegible]	Hayır
5	3.5.2021 21:20:44	[illegible]	Evet
6	3.5.2021 21:20:50	[illegible]	Evet
7	3.5.2021 21:20:50	[illegible]	Evet
8	3.5.2021 21:20:50	[illegible]	Evet
9	3.5.2021 21:20:50	[illegible]	Evet
10	3.5.2021 21:20:52	[illegible]	Evet
11	3.5.2021 21:20:52	[illegible]	Evet
12	3.5.2021 21:20:53	[illegible]	Evet
13	3.5.2021 21:20:56	[illegible]	Hayır
14	3.5.2021 21:20:59	[illegible]	Evet
15	3.5.2021 21:21:04	[illegible]	Evet
16	3.5.2021 21:21:05	[illegible]	Hayır
17	3.5.2021 21:21:05	[illegible]	Evet
18	3.5.2021 21:21:06	[illegible]	Evet
19	3.5.2021 21:21:07	[illegible]	Evet
20	3.5.2021 21:21:08	[illegible]	Evet
21	3.5.2021 21:21:08	[illegible]	Evet
22	3.5.2021 21:21:09	[illegible]	Evet

Form Yanıtları 1

2.5.5. *Demographic Data Form*

The researcher created it based on a literature review to collect demographic information about the participants. The participants were asked about their age, gender, grade level, computer knowledge, ML experiences, and other relevant information.

2.6. Data Analysis

The achievement test (as a pretest and a posttest), Moodle activity data, Demographic Data Form, Course Attendance Form, and Online Student Engagement Scale data were coded using Microsoft Excel. Then, the analysis was conducted using SPSS 21. For the 38 questions in the achievement test, 1 point was given for each correct answer and 0 for each wrong answer, both in the pre-test and post-test. Thus, participants could get a total of 38 points. The answers to the Online Student Engagement Scale were coded between 1 and 5, depending on the scale items. Participants could get a maximum of 95 points on the 19-item scale. Then, the scores were converted into a 100-point system to ensure equivalent evaluations.

Through a Microsoft Excel document obtained from the system, Moodle activity data were coded as 1 for completed activities and 0 for incomplete activities. The engagement data were gathered using the Course Attendance Form for 8 weeks, and coding was done by giving 1 point for those who attended the lesson and 0 points for those who did not. Participants who attended all lessons for 8 weeks received 8 points.

After coding the achievement test as the pretest and posttest, the data were analyzed using SPSS according to the research questions, employing t-test, ANOVA, and multiple linear regression (MLR) analysis procedures.

2.7. Validity And Reliability

The analyses were designed to produce appropriate results while adhering to ethical guidelines. At this point, opinions from experts in the field were solicited at regular intervals.

2.7.1. Item Analsysis of The Achievement Test (Pretest and Posttest)

The achievement test, used as the pretest and posttest for the ML training, was pilot tested, and item difficulty (p) and item discrimination (r) were analyzed. There were 30 questions with high item discrimination ($r > 0.40$), 6 questions with item discrimination between 0.25 and 0.39, and 2 questions with item discrimination of $r < 0.20$. Upon examination, it was determined with the experts that the question "Which of the following is not one of the project types in ML for Kids?" lacked meaningful expression. Consequently, the question was revised to "Which of the following is not among the types of projects prepared in ML for Kids?"

Item difficulty was 0.75 for 32 of the questions. This was attributed to ML being a new subject, and the participants' lack of knowledge about it, as concluded by experts in the field. As a result, the average item difficulty (p) was found to be 0.72. The developed achievement test was determined to have suitable validity and reliability based on the analysis.

2.7.2. Reliability of The Online Student Engagement Scale

Since the factor structure of the original scale was deemed appropriate by experts in the study by Polat et al. (2022), confirmatory factor analysis (CFA) was performed to test the construct validity of the scale. The Cronbach alpha internal consistency coefficient ranged from $\alpha = 0.77$ to $\alpha = 0.87$ in the study, which included 254 university students. Separate t-tests were performed to examine the significance levels of the mean values. Since this value was 0.00 ($p > 0.05$), it was accepted as a good fit. The $\chi 2$ value was 273.844, and the standard deviation (SD) was 142. The

AGFI was found to be acceptable, and all other indices indicated a good fit. Thus, it was confirmed that the model had four factors. The t-value for all of the items was greater than +1.96 and was significant ($p < 0.05$). When the data collected in the study were analyzed, the Cronbach alpha internal consistency coefficient was determined to be 0.95. As a result, it was determined that using the scale was appropriate.

3. Result

The collected data were analyzed, and the findings were determined after the ML training. The research question, "Is the ML training for preservice teachers in the STEM field effective?" was examined and analyzed. The results were derived from the preservice teachers' responses to the achievement test before and after the ML training. The achievement test, used as both the pretest and posttest, consisted of 38 multiple-choice questions, each with 5 choices. Scores on the scale ranged from a minimum of 0 to a maximum of 100.

Table 4

Descriptive Statistics for The Achievement Test (Pretest and Posttest)

	N	M	Min.	Max.	Skewness	Kurtosis
Pretest	74	59.51	26.40	84.48	-0.09	-0.08
Posttest	74	76.02	50.16	92.40	-0.46	-0.05

According to the pretest and posttest descriptive statistics of the preservice teachers shown in Table 3, 74 people participated in both the pretest and the posttest. The average, minimum, maximum, skewness, and kurtosis values of the pretest were 59.51, 25.40, 84.48, -0.09, and -0.08, respectively. The average, minimum, maximum, skewness, and kurtosis values of the posttest were 76.02, 50.16, 92.40, -0.46, and -0.05, respectively.

Table 5

The T-Test Results for The Achievement Test (Pretest and Posttest)

	N	M	SD	t	P
Pretest	74	59.51	11.66	11.58	0.000
Posttest	74	76.02	9.35		

When Table 5 is examined, a significant difference is seen between the pretest and posttest mean scores of the preservice teachers in favor of the posttest ($t = 11.58$, $p < 0.05$). To test the magnitude of the resulting difference, the effect size was calculated. The effect size was $d = 1.56$. The analysis revealed a significant increase in the mean posttest score compared to the pretest score, suggesting that the machine learning training positively impacted the preservice teachers' learning of ML.

Analyses were conducted for the research question, "Is there a significant difference between the pretest and posttest achievement test scores depending on departments?" The results were derived from the preservice teachers' responses to the achievement test before and after the ML training and the Demographic Data Form. The achievement test, used as the pretest and posttest, consisted of 38 multiple-choice questions, each with 5 choices. Scores ranged from a maximum of 100 to a minimum of 0 on the scale. The departments were determined based on the demographic data obtained prior to the training.

Table 6

Descriptive Statistics for The Achievement Test (Pretest and Posttest) By Departments

Department		N	M	Min.	Max.	SD
Science Education	Pretest	21	58.83	44.88	79.20	10.50
	Posttest	21	73.54	52.80	87.12	8.69
Mathematics Education	Pretest	20	57.82	34.32	84.48	12.70
	Posttest	20	76.69	55.44	92.40	10.47
CEIT	Pretest	33	60.96	26.40	84.84	11.90
	Posttest	33	77.20	50.16	92.40	9.03

According to Table 6, the mean total score of the preservice teachers who took the pretest was 58.83 in the science education department, 57.82 in the mathematics education department, and 60.96 in the CEIT department. For the posttest, the mean total score was 73.54 in the science education department, 76.69 in the mathematics education department, and 77.20 in the CEIT department. ANOVA was performed to determine whether the pretest and posttest (achievement test) scores differed depending on the departments, and the results are given in Table 7.

Table 7

ANOVA Results for The Achievement Test (Pretest and Posttest) By Departments

		Sum of Squares	df	Mean Square	F	p
Pretest	Between departments	136.36	2	68.18	0.49	0.61
	Within departments	9804.93	71	138.10		
	Total	9941.29	73			
Posttest	Between departments	183.84	2	91.92	1.05	0.35
	Within departments	6200.03	71	87.32		
	Total	6383.87	73			

According to Table 7, there was no significant difference between the pretest and posttest total scores depending on the departments in the training taken by the preservice teachers in the departments of science education, mathematics education, and CEIT (pretest $F = 0.49$, $p > 0.05$ and posttest $F = 1.05$, $p > 0.05$). When the mean scores were examined, it was found that the biggest difference was among the preservice teachers in the mathematics education department.

Analyses were conducted for the research question, "Is there a significant difference between the pretest and posttest achievement test scores depending on practical and theoretical questions?" The results were derived from the preservice teachers' responses to the achievement test before and after the ML training. The achievement test, used as both the pretest and posttest, consisted of 38 multiple-choice questions, each with 5 choices. Scores ranged from a minimum of 0 to a maximum of 100. The analysis was conducted by separating the topics in the specifications table as practical and theoretical questions. There were 13 practical questions and 25 theoretical questions. Since the total score of the practical and theoretical questions was calculated as 100, each correct answer was multiplied by 2.64. Thus, the highest score to be obtained from the practical questions was 34, and the highest score from the theoretical questions was 66. T-test analysis was performed to determine whether there was a difference between the pretest and posttest (achievement test) scores of the 74 preservice teachers who took part in the study depending on practical and theoretical questions.

Table 8

Descriptive Statistics for The Achievement Test (Pretest and Posttest) Practice Questions

	N	M	Min.	Max.	Skewness	Kurtosis
Pretest Practice Questions	74	40.56	15.84	55.44	-0.56	0.08
Posttest Practice Questions	74	53.16	34.32	63.36	-0.99	0.82

According to Table 8, based on the descriptive statistics of the pretest and posttest (achievement test) practical questions, the average, minimum, maximum, skewness, and kurtosis of the pretest practical questions were 40.56, 15.84, 55.44, -0.56, and 0.08, respectively. The average, minimum, maximum, skewness, and kurtosis values of the posttest practical questions were 53.16, 34.32, 63.36, -0.99, and 0.82, respectively.

Table 9

T-Test Results for The Achievement Test (Pretest and Posttest) Practice Questions

	N	M	SD	T	P
Pretest Practice Questions	74	40.56	8.46	124.53	0.000
Posttest Practice Questions	74	53.16	6.25		

Table 9 reveals a significant difference in the mean scores of the pretest and posttest practical questions ($t = 124.53$, $p < 0.05$). The effect size was found to be $d = 1.69$. As a result, the average score for the practical questions in the posttest was higher than the average score for the practical questions in the pretest.

Table 10

Descriptive Statistics for The Achievement Test (Pretest and Posttest) Theoretical Questions

	N	M	Min.	Max.	Skewness	Kurtosis
Pretest Theoretical Questions	74	18.94	2.64	31.68	-0.09	0.51
Posttest Theoretical Questions	74	22.87	10.56	34.32	-0.26	0.25

According to Table 10, based on the descriptive statistics of the pretest and posttest (achievement test) theoretical questions, the average, minimum, maximum, skewness, and kurtosis of the pretest theoretical questions were 18.94, 2.64, 31.68, -0.09, and 0.51, respectively. The average, minimum, maximum, skewness, and kurtosis values of the posttest theoretical questions were 22.87, 10.56, 34.32, -0.26, and 0.25, respectively.

Table 11

T-Test Results for The Achievement Test (Pretest and Posttest) Theoretical Questions

	N	M	SD	T	p
Pretest Theoretical Questions	74	18.94	5.32	5.15	0.000
Posttest Theoretical Questions	74	22.87	5.30		

Table 11 reveals a significant difference in the mean scores of the pretest and posttest theoretical questions ($t = 5.15$, $p < 0.05$). The effect size was found to be $d = 0.74$. As a result, the average score for the theoretical questions in the posttest was higher than the average score for the theoretical questions in the pretest.

Analyses were carried out for the research question, "To what extent does engagement in ML training given to preservice teachers in the STEM field predict achievement?" The data for this research question came from the Online Student Engagement Scale, Moodle activity, and Course Attendance Form. MLR analysis was used to determine how well each variable predicted the posttest (achievement test).

The Online Student Engagement Scale had 19 5-point Likert-type questions. These were: (1) Does not define me at all, (2) Does not define me, (3) Undecided, (4) Defines me, and (5) Defines me completely. The highest possible score was 100. Moodle activity data included data from 29 activities, consisting of course content and course records shared with the preservice teachers during the training. The total score from Moodle activity data was 100.

The Course Attendance Form was completed online in each course to determine the status of engagement in the weekly classes over the course of eight weeks. The analysis was carried out by calculating the average of the Course Attendance Form. The highest possible score was 100.

Table 12

Descriptive Statistics Results for The Moodle Activity Data, Online Student Engagement Scale, And Course Attendance Form

	N	M	SD	Min.	Max.	Skewness	Kurtosis
Moodle Activity Data	74	82.36	11.50	52	100	0.05	-0.40
Online Student Engagement Scale	74	73.62	15.50	23	100	-1.21	1.66
Course Attendance Form	74	94.39	8.22	75	100	-1.18	0.22

According to Table 12, 74 people participated in Moodle activities. The average, skewness, kurtosis, minimum, maximum, and standard deviation values of the data were 82.36, 0.05, -0.40, 52, 100, and 11.50, respectively. A total of 74 people filled out the Online Student Engagement Scale. The average, skewness, kurtosis, minimum, maximum, and standard deviation values of the scale were 73.62, -1.21, 1.66, 23, 100, and 15.50, respectively. Finally, 74 people were recorded in the Course Attendance Form. The average, skewness, kurtosis, minimum, maximum, and standard deviation values of the form were 94.39, -1.18, 0.22, 75, 100, and 8.22, respectively.

Table 13

Results for Relationships Between Variables

	Posttest	Moodle Activity Data	Online Student Engagement Scale
Moodle Activity Data	0.15	-	-
Online Student Engagement Scale	0.34	0.04	-
Course Attendance Form	0.38**	0.33**	0.30**

$^{**}p < 0.01$, $^{*}p < 0.05$

The posttest was found to have a significant relationship with the Online Student Engagement Scale and the Course Attendance Form, as shown in Table 13.

Table 14

MLR Results for The Variables Affecting Posttest (Achievement Test) Scores

R	R^2	Adjusted R^2	F	P
0.45	0.20	0.17	6.242	0.001

Table 14 shows the MLR analysis conducted to test the degree to which the Moodle activity, Online Student Engagement Scale, and Course Attendance Form data of the preservice teachers studying in the mathematics education, science education, and CEIT departments could predict the posttest scores.

4. Discussion and Conclusion

This study was prepared to plan, implement, and evaluate ML training for preservice teachers in the STEM field to contribute to the development of the computer science field's future workforce. It was found that there was a significant difference between the mean pretest and posttest (achievement test) scores in favor of the posttest. As a result, the provided training has had a positive impact on preservice teachers' ML. The application of ML in educational practices is a relatively new topic in the literature. For this reason, no study could be found in the literature that provides ML training to preservice teachers using ML for Kids. However, there are studies in the literature on training given to elementary and middle school students (Martin et al., 2024). Chklovski et al. (2019) trained 3rd and 8th-grade students and found them to be increasingly interested. In ML training, homework, discussion platforms, and guest educator seminars have been observed to be effective in providing different perspectives beyond learning new knowledge. As part of the study, preservice teachers studying in the STEM fields in a faculty of education were trained.

There was a significant difference between the practical and theoretical questions in the pretest and the posttest (achievement test). Hitron et al. (2019), who reached the same conclusion through a sample with different characteristics, found a significant difference between the pretest and posttest in the training offered to 10–13-year-old children.

There was no significant difference between the pretest and posttest scores depending on the departments. When the averages were examined, it was found that the highest difference was among the preservice teachers who studied in the mathematics education department. In the current study, the sample included preservice teachers in the STEM field in a faculty of education. Ahmad et al. (2020) conducted a study on undergraduate students in the department of physics and stated that academic departments were an important predictor in measuring students' performance. Similarly, Buyruk and Korkmaz (2016), who conducted a study on a sample having the same characteristics as the sample of the current study, stated that the highest levels of awareness of the STEM approach were in science education, CEIT, and mathematics education students, in the order given.

How the posttest predicted engagement data was examined, revealing that the posttest was significant in predicting the Online Student Engagement Scale and Moodle activity data. Yildiz (2014), who examined Moodle data in a similar manner, stated that ML predictions reduced error rates. Gok (2017), who conducted a study on a sample with different characteristics, stated that the success rate was high when regression and classification methods were used to predict the overall success of 6th, 7th, and 8th-grade students.

In future studies, training can be implemented face-to-face. The spread of ML can be aided by developing resources that are accessible to all students and teachers. At this point, a variety of platforms may also be preferred for diversity instead of just relying on the ML for Kids platform. Language support on an ML platform is extremely useful during the learning phase. Moreover, because there is no ML platform specific to Turkey, developing an ML platform can contribute to the literature. The research is limited to university students. Additionally, providing education online due to pandemic conditions is among other limitations.

5. References

Ahmad, S. A., Hussain, I., Ahmad, R., & Din, M. N. U. (2020). Performance based prediction of the students in the physics subject using traditional and machine learning approach at higher education level. *International Journal of Innovation in Teaching and Learning (IJITL)*, *6*(1), 174-190. https://doi.org/10.35993/ijitl.v6i1.997

Buyruk, B., & Korkmaz, O. (2016). Teacher candidates STEM awareness levels. *Online Submission*, *4*(1), 272-279. https://dergipark.org.tr/en/download/article-file/777013

Buyukozturk, S., Cakmak, E. K., Akgun, O. E., Karadeniz, S., & Demirel, F. (2008). *Scientific research methods.* Pegem academy. https://avesis.medeniyet.edu.tr/yayin/d93be4a3-e3bb-42ae-8635-96946af53ee9/bilimsel-arastirma-yontemleri

Bybee, R. W. (2010). What is STEM education?. *Science*, *329*(5995), 996. https://www.science.org/doi/10.1126/science.1194998

Cevik, K. K., & Kayakus, M. (2020). Prediction of solution time of user requests delivered to the information technologies department via machine learning. *Journal of Engineering Sciences and Design*, *8*(3), 728-739. https://doi.org/10.21923/jesd.722323

Chklovski, T., Jung, R., Fofang, J. B., Gonzales, P., Hub, B. T., & La Paz, B. (2019). Implementing a 15-week AI-education program with under-resourced families across 13 global communities. *In International joint conference on artificial intelligence.*

Chung, C. J., & Shamir, L. (2020). Introducing machine learning with scratch and robots as a pilot program for k-12 computer science education. *Science Education*, *6*, 7. https://people.cs.ksu.edu/~lshamir/publications/ICFL2020.pdf

Demirkaya, H., & Sarpel, E. (2018). In the training and development applications virtual reality from the new generation computer technologies, cluster computing and artificial intelligence. *Karadeniz Uluslararasi Bilimsel Dergi,* (40), 231-245. https://doi.org/10.17498/kdeniz.460145

Dixson, M. D. (2015). Measuring student engagement in the online course: The online student engagement scale (OSE). *Online Learning, 19*(4), n4. https://files.eric.ed.gov/fulltext/EJ1079585.pdf

Gok, M. (2017). Predicting academic achievement with machine learning methods. *Gazi University Journal of Science and Technology Part C: Design and Technology, 5*(3), 139-148. https://dergipark.org.tr/en/download/article-file/341510

Hitron, T., Orlev, Y., Wald, I., Shamir, A., Erel, H., & Zuckerman, O. (2019, May). *Can children understand machine learning concepts? the effect of uncovering black boxes.* In Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems. England. https://doi.org/10.1145/3290605.3300645

Karasar, N. (2017). *Preparing report in research.* Nobel publication distribution.

Kayahan, S. (2018). Artificial intelligence in educational applications: A review. *University of Trakya, 201*, 20.

Kim, K., & Kwon, K. (2024). A systematic review of the evaluation in K-12 artificial intelligence education from 2013 to 2022. *Interactive Learning Environments*, 1-29. https://doi.org/10.1080/10494820.2024.2335499

Lane, D. (2018) Explaining artificial intelligence. *Hello World, 4*, 44-46.

Martin, F., Vahedian Movahed, S., Dimino, J., Farrell, A., Irankhah, E., Ghosh, S., ... & Narain, S. (2024). Perception, Trust, Attitudes, and Models: Introducing Children to AI and Machine Learning with Five Software Exhibits. *In Proceedings of the 55th ACM Technical Symposium on Computer Science Education V.* 2 (pp. 1740-1741). https://doi.org/10.1145/3626253.3635512

Mduma, N., Kalegele, K., & Machuve, D. (2019). A survey of machine learning approaches and techniques for student dropout prediction. *Data Science Journal, 18*, 1–10. https://doi.org/10.5334/dsj-2019-014

Murphy, R. F. (2019). *Artificial intelligence applications to support k–12 teachers and teaching*. RAND Corp.,USA. https://www.rand.org/pubs/perspectives/PE315.html

Nabiyev, V. V. (2012). *Artificial intelligence: human-computer interaction. Seckin Publishing.*

Nafea, I. T. (2018). Machine learning in educational technology, H. Farhadi (Ed.), *Machine learning - advanced techniques and emerging applications* (s. 175-183). Intech Open. https://www.jstor.org/stable/pdf/resrep19907.pdf

Park, D., Ahn, J., Jang, J., Yu, W., Kim, W., Bae, Y., & Yoo, I. (2020). The development of software teaching-learning model based on machine learning platform. *Journal of The Korean Association of Information Education,* 24(1), 49-57. https://doi.org/10.14352/jkaie.2020.24.1.49

Peters, L. (2019). *An educational programming environment for teaching the principles of machine learning using lego mindstorms* [Unpublished doctoral dissertation]. University of Applied Sciences Leipzig.

Polat, E., Hopcan, S., & Arslantaş, T. K. (2022). Çevrimiçi öğrenci bağlılık ölçeğinin Türkçe'ye uyarlanması: geçerlik ve güvenirlik çalışması. *Eğitim Teknolojisi Kuram ve Uygulama*, *12*(1), 41-56.

Priya, S., Bhadra, S., Chimalakonda, S., & Venigalla, A. S. M. (2024). ML-Quest: a game for introducing machine learning concepts to K-12 students. *Interactive Learning Environments*, *32*(1), 229-244. https://doi.org/10.1080/10494820.2022.2084115

Quiroz, P., & Gutierrez, F. J. (2024). Scratch-NB: A Scratch Extension for Introducing K-12 Learners to Supervised Machine Learning. In *Proceedings of the 55th ACM Technical Symposium on Computer Science Education,* V. 1 (pp. 1077-1083). https://doi.org/10.1145/3626252.3630920

Reyes, A. A., Elkin, C., Niyaz, Q., Yang, X., Paheding, S., & Devabhaktuni, V. K. (2020, August). A preliminary work on visualization-based education tool for high school machine learning education. In *2020 IEEE Integrated STEM Education Conference (ISEC), IEEE*. https://ieeexplore.ieee.org/abstract/document/9280629

Sakulkueakulsuk, B., Witoon, S., Ngarmkajornwiwat, P., Pataranutaporn, P., Surareungchai, W., Pataranutaporn, P., & Subsoontorn, P. (2018). Kids making AI: integrating machine learning, gamification, and social context in STEM education. In *2018 IEEE International Conference on Teaching, Assessment, and Learning for Engineering*. https://ieeexplore.ieee.org/abstract/document/8615249

Senocak, D. (2020). Artificial intelligence in open and distance learning environments: Its opportunities and the concerns raised. *Journal of Open Educational Practices and Research, 6*(3), 56-78. https://dergipark.org.tr/en/download/article-file/1191727

Tataj, X., & Kola, M. (2021). Education policies in line with the latest developments in the field of artificial intelligence: case of Albania. *Human & Human, 8*(27), 101-116. https://doi.org/10.29224/insanveinsan.818263

Tektas, M., Tektas, N., Onat, N., Gokmen, G., Kocyigit, G., & Akinci, T. Ç. (2010). Preparation of web based artificial intelligence techniques training simulators. *Marmara University BAP Commission Presidency Project Report*, 76-90. https://docplayer.biz.tr/631443-Web-tabanli-yapay-zeka-tekn-kler-proje-no-fen-e-050608-138.html

Tseng, T., Davidson, M. J., Morales-Navarro, L., Chen, J. K., Delaney, V., Leibowitz, M., ... & Shapiro, R. B. (2024). Co-ML: Collaborative machine learning model building for developing dataset design practices. *ACM Transactions on Computing Education, 24* (2), 1-37. https://doi.org/10.1145/3641552

Wang, H. H., Moore, T. J., Roehrig, G. H., & Park, M. S. (2011). STEM integration: Teacher perceptions and practice. *Journal of Pre-College Engineering Education Research (JPEER), 1*(2), 2. https://doi.org/10.5703/1288284314636

Yildiz, O. (2014). *Evaluation of the performance of distance education students with machine learning*. [Unpublished doctoral dissertation]. Istanbul University.

Zhu, K. (2019). *An educational approach to machine learning with mobile applications* [Unpublished doctoral dissertation]. Massachusetts Institute of Technology.