A systematic review of teaching and learning machine learning in K-12 education

Ismaila Temitayo Sanusi[1] · Solomon Sunday Oyelere[2] ·
Henriikka Vartiainen[3] · Jarkko Suhonen[1] · Markku Tukiainen[1]

Abstract

The increasing attention to Machine Learning (ML) in K-12 levels and studies exploring a different aspect of research on K-12 ML has necessitated the need to synthesize this existing research. This study systematically reviewed how research on ML teaching and learning in K-12 has fared, including the current area of focus, and the gaps that need to be addressed in the literature in future studies. We reviewed 43 conference and journal articles to analyze specific focus areas of ML learning and teaching in K-12 from four perspectives as derived from the data: curriculum development, technology development, pedagogical development, and teacher training/professional development. The findings of our study reveal that (a) additional ML resources are needed for kindergarten to middle school and informal settings, (b) further studies need to be conducted on how ML can be integrated into subject domains other than computing, (c) most of the studies focus on pedagogical development with a dearth of teacher professional development programs, and (d) more evidence of societal and ethical implications of ML should be considered in future research. While this study recognizes the present gaps and direction for future research, these findings provide insight for educators, practitioners, instructional designers, and researchers into K-12 ML research trends to advance the quality of the emerging field.

Keywords Machine learning · Artificial intelligence · K-12 · Systematic review

1 Introduction

Popular interest in artificial intelligence (AI) has increased incredibly in recent times. Especially, machine learning (ML), an essential subset of AI that has become the new engine that revolutionizes practices of knowledge discovery (Lin et al.,

✉ Ismaila Temitayo Sanusi
  ismaila.sanusi@uef.fi

Extended author information available on the last page of the article

2020). The benefits attributed to introducing K-12 learner to the intricacies and inner workings of machine learning has led to increased interest among educational stakeholders (Touretzky et al., 2019). Such that ML education in K-12 is considered relevant for future generations who must live in the ML era. As machine learning is becoming a commonplace feature of people's everyday lives, early introduction to the fundamental processes of ML can ease children's understanding of the world around them (Hitron et al., 2018). According to Lin et al. (2020) understanding how machines learn is critical for children to develop useful mental models for exploring AI and smart devices that they now frequently interact with. Introducing the basics of ML will also motivate the next generation of AI researchers and software developers (Touretzky et al., 2019).

Though teaching the fundamentals of ML concepts and techniques is common in higher education institutions (Ho & Scadding, 2019; Sulmont et al., 2019), there have been attempts in recent times to introduce the teaching of the emerging concept in K-12. For instance, Sabuncuoglu (2020) designed a year curriculum to teach AI for middle school, and Opel et al. (2019) also developed teaching materials on AI using a simulation game. Furthermore, Scheidt and Pulver, (2019) presented Any-Cubes, a prototype toy with which children can intuitively and playfully explore to understand machine learning while Lin et al. (2020) designed Zhorai, a conversational agent for Children to explore ML concepts. Studies keep growing on either curriculum design, materials, tools, or platforms for machine learning in schools across countries (Lin et al., 2020; Williams et al., 2019a). This presents a challenge as these launched initiatives or projects seem to be scattered which makes it troublesome to get a synopsis of the extant works and focus of studies in ML education. Besides, K-12 curricula do not precisely present the AI concept (Ho & Scadding, 2019) and only eleven countries have government-endorsed K-12 AI curricula (Miao & Holmes, 2022).

Recent studies have attempted to review literature in this field. The studies include Marques et al., (2020) which conducts a systematic mapping study focusing on instructional units accessible for introducing machine learning for K-12 levels. Zhou et al. (2020) also carried out an exploratory review of the existing AI learning tools and curriculum to analyze how designs contributed to building K-12 AI literacy while Giannakos et al. (2020) presented a review of games to teach AI and ML. However, this study differs from earlier research as it presents a systematic review of the focus area of research on ML teaching and learning in K-12, specifically as it relates to curriculum, technology, pedagogy, and professional development. It is important to understand the tools available, the curriculum designed, the kind of pedagogies that exist, and training needs available for the teachers in the field of AI/ML. In addition, it is important to uncover the potential issues that are related to this new area in order to nurture the emerging field. The issues are reported as limitations of earlier research which could either be human, data, or technology related. The importance of this research stems from being a contribution to the field of machine learning education as it systematically synthesizes research conducted on teaching and learning machine learning at K-12 levels. The recent call to incorporate ML ideas and concepts in school curriculum (Sanusi et al., 2022b; Touretzky et al., 2019) neccesitated the need for this research. This emerging area of research

currently generate interest globally among researchers, practitioners and educators. Consequently, several initiatives to promote ML within K-12 levels keep growing. The increasing effort in the research area requires that we synthesize existing studies to understand how the concept has been introduced in the past and identified areas for future research. According to Papamitsiou and Economides (2014), 'the value of any single study is derived from how it fits with and expands previous work.' As a result, summing up empirical evidence from prior research would contribute to the understanding of the domain. Machine learning education in K-12 level is a relatively new area of research interest which requires more study to ensure its effective integration into school system. Equally important is putting together results of exisiting works to understand how research that introduces ML to K-12 students has fared, considering the current area of research focus as well as revealing gaps in literature. Identifying the highlighted areas will inform the scientific community of approaches, resources and opportunities that exist for ML implementation in schools as well as further research possibilities.

This work can assist teachers and instructional designers as they try to plan and design ML resources and integrate ML technologies into their teaching practice (Chiu & Chai, 2020). The present study systematically reviews the focus area of existing literature and identify gaps including future directions in K-12 ML research. Although ML education in K-12 is an emerging research area, there exists enough work to perform a review and generate insights.

The research objectives are:

RO1: To classify the present focus area of research on ML teaching and learning in K-12

RO2: To investigate the issues and suggest future direction of research on ML teaching and learning in K-12

This article is organized as follows: Sect. 2 discusses a review of related works – it includes machine learning in K-12 education and previous work. Section 3 evinces the methodology adopted which includes details on planning, conducting, and reporting of the review. The analysis of the findings as obtained from the systematic review process as provided in Sect. 4. The findings were discussed in Sect. 5 while the study was finally concluded with suggestions and review limitations in Sect. 6.

2 Background

2.1 Machine Learning Education

The machine learning concept originated in the 1950s and it was studied as a separate field in the 1990s (Michalski et al., 2013). Education research on machine learning remains nascent, in large part because machine learning is a relatively new subject in university curricula (Sulmont et al., 2019). Georgiopoulos et al. (2009) stated that courses on ML have existed for decades at many academic institutions. Machine

Learning has considered a sub-field of computer science and the students in ML courses are likely to be in computer science or a closely related field such as data science (Heys, 2018). Machine learning is currently being taught to university students during their second and third years in many courses including those that are outside of computer science (Rattadilok and Roadknight, 2018). Machine learning technology is enabling a paradigm shift in problem-solving from analytical to a powerful data-driven approach, through computer programs that learn models from training data and predict results from new data (Huang & Ma, 2018). Studies on teaching machine learning are not overwhelmingly reported as machine learning application literature. Of the relatively few literatures, Ho and Scadding (2019) reported that the challenge frequently encountered by teachers who introduce ML-related content is the difficulty in teaching technologically related subjects to learners who may not have much interest in technology. In most cases, teaching machine learning centers around two key components: ensuring the understanding of a machine learning technique and the usage of such machine learning technique. The understanding part usually involves giving an introduction about a particular machine learning technique, and the application of that particular machine learning technique on sample data sets using a particular machine learning software tool.

## 2.2 Reviews on K-12 Machine Learning

Even though no systematic reviews have been identified that explicitly examine the current focus area of K-12 ML (as it relates to curriculum, technology, pedagogy, and professional development) including issues that require further research, we have identified seven literature reviews on K-12 AI/ML education. These reviews give an overview of research studies on instructional units and tools available for teaching ML, games used within the field, and definition of AI literacy. Using 33 studies (2009–2019), Marques et al. (2020) performed a systematic mapping study (SMS) to analyze the instructional units presently obtainable to teach ML in K-12. They identified 30 instructional units (IU) with a prevalent focus on the basics of ML and neural networks. With regards to the complexity of ML concepts, numerous instructional units focus only on the utmost accessible processes, which include data management or model learning and testing on an abstract level black boxing some of the fundamental ML processes. The key finding of the paper is the aggregating of IU features for ML education across K-12 levels concerning content, context, and the scrutiny of how they were established and appraised. The authors observed that several IUs mostly focus on the educational stage from elementary to high school and present ML concepts on an abstract level as supplementary units varying from an hour introductory workshop to semester-long courses. Marques et al. (2020) however indicated a lack of organized demonstration of the IUs and how they were created and assessed a significant number of the articles reviewed are non-scientific. The second review was conducted by Giannakos et al. (2020) and addressed the use of games within the field. Giannakos et al. (2020) presented an overview of the relevant research papers on games with K-12 ML as well as showcased how different games provide a unique opportunity to teach several different concepts and topics in

AI and ML. The comprehensive review provided a guide to different stakeholders to explore and put into practice the games that meet their needs as well as presented open RQs in this educational field. After reviewing the studies, they identified 17 games/projects. The third review by Zhou et al. (2020) analyzed recently published articles concerning AI Learning Experiences for K-12. Zhou et al. (2020) conducted an exploratory review of AI for K-12 education (AI4K12) literature and tools to distill a design framework to inform the development of AI learning experiences for K-12. They identified future opportunities and K-12-specific design guidelines, to support researchers, designers, and educators in creating K-12 AI learning experiences. The identified design guidelines include those related to student engagement, built-in scaffolding, teacher and parent involvement, equity diversity and inclusion, and integrated AI/core curricula.

Sanusi and Oyelere (2020) conducted the fourth review using a narrative approach to reveal pedagogies associated with teaching K-12 ML. The study suggested that learner-centered pedagogies approach such as participatory learning, design-oriented learning, and active learning could be appropriate for teaching ML in K-12. It was argued that this approach allows the students to be active in learning. With 24 literature (2010–2020), Gresse von Wangenheim et al. (2022) present a systematic mapping of tools that support the teaching of ML at K-12 as well as analyze the identified tools per their educational features, support for ML models development including how the tools have been designed and appraised. They come across 16 tools aimed at students typically as part of short-duration extracurricular activities with the result that shows the tools can successfully leverage students' comprehension of ML. The sixth review conducted by Tedre et al. (2021) was a scoping review that charts the evolving paths in educational practice, theory, and technology connected with ML at the K-12 levels. The paper centers on the main features of the paradigm shift that will be required to successfully integrate ML into K-12 computing curricula. The last review conducted by Ng et al. (2021) reviewed 30 articles from 2016 to 2021 and investigated how scholars describe AI literacy including how it can be learned and the ethical concerns. Ng et al. (2021) study identified a variety of definitions which was mostly based on a different type of literacies useful to define skill sets in other disciplines. This study extends previous reviews on K-12 ML research by placing emphasis on mapping the existing literature based on their area of focus to determine the present state of knowledge and aggregate published articles to provide possible directions about ML in K-12 education.

3 Methodology

Synthesizing existing research in K-12 ML is important with the increased interest in the subject in recent times. More research must be carried out to detect new open problems and trends, and further, enrich the knowledge base. As a result, this study was designed using a systematic literature review method to understand the development of ML education given the guiding principle by Kitchenham and Charters (2007). Systematic literature review method comprised of three main stages based on guidelines shown in Fig. 1. No single source finds all the primary
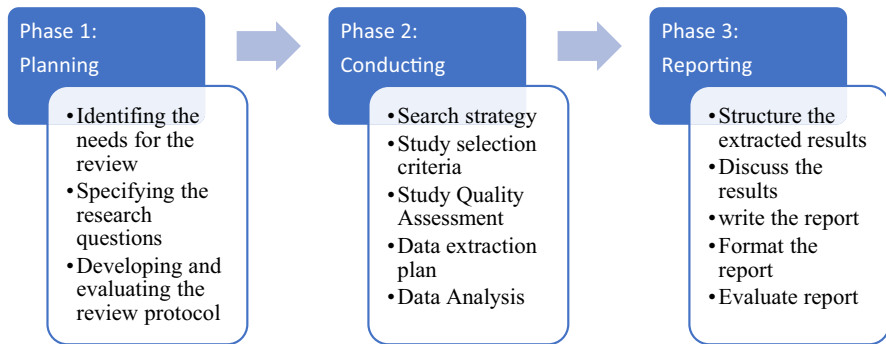
Fig. 1 Systematic literature review process (Kitchenham & Charters, 2007)

studies, this study retrieved articles from six publication databases as seen in Table 1. The six publication databases were selected since they contain publications focusing on computing fields and engineering education which is related to the focus of the study. The search results from publication databases were downloaded in comma-separated value (CSV) format. Those unable to be downloaded were done manually with copy and pasting from the databases into the Microsoft word text editor. Excel worksheets were utilized to compute the data obtained in an organized manner while the duplicates were manually sorted out.

3.1 Planning the review

3.1.1 Identifying the need for the review

Based on the above-mentioned procedure, the number one step is to strategize and check that the review conditions are well planned. The plan of a review begins by recognizing the main aim connected to teaching machine learning at K-12 levels literature.

Table 1 Search structure

Academic Databases	Search design
Scopus	Title
Web of Science (WoS)	Topic
IEEE explore	Abstract
ACM	Abstract
Science Direct	Alternative search
Springer Link	Title

3.1.2 Specifying the RQs

Our paper included existing research conducted in the past without limitations to a specific year range. We framed the research questions below to categorize the papers for data extraction and analysis:

> RQ1: What is the current focus area of research on ML teaching and learning in K-12?
> We analyzed the selected articles given the present focus area of literature to understand the existing research in the emerging field. The focus areas derived from the analyzed data were categorized under four headings: curriculum development, technology development, pedagogical development, and teacher training/professional development.
> RQ2: What are the issues that should be investigated and future directions of research on ML teaching and learning in K-12?

This study investigates the limitations of the research conducted to demystify ML to K–12 students. It equally summarizes and suggests the shortcomings discovered in the literature, which is useful for future direction and research.

3.1.3 Developing the protocol for review

The approach employed in this review was founded on the protocol for review. Specifying the methods to be adopted in the review in advance help lessen the risk of unintended errors. While planning the review, we applied informal and formal searches to find the research objectives and glean relevant information required for the study. The initial result is presented as previous studies in the background section. This assists in creating research questions to guide the review process.

3.2 Conducting the review

This section explicates how the review is been conducted. It includes the search strategy, selection criteria, study quality assessment, data extraction plan, and analysis of data.

3.2.1 Search strategies

This study formulated the search strategy resulting from research objectives. To limit the number of irrelevant papers, keywords were recognized, search strings were created, the search structure was framed, and the process of the search was finally carried out. Search strings were generated to achieve a result based on the study's research objectives.

3.2.2 Search keywords

The set of keywords was generated to gather studies about machine learning for K-12. We collected existing keyword-search approaches from the relevant literature reviews on teaching ML at K-12 levels (Marques et al., 2020), a review of games for AI and ML education (Giannakos et al., 2020), and review of general AI literacy competencies (Zhou et al., 2020). Specific and same keywords were not used across all the digital publication channels because the keywords are longer and not possible on some databases. Table 1 evince the search structure and Table 2 shows the protocol executed for each database.

The search structure in Table 1 evince the academic databases and the search design – where the search was conducted in each database. In Scopus and WoS, title and topic search include the article title, abstract, and keyword. The abstract was searched in IEEE explore and ACM. For science direct, the alternative search means – "find articles with these terms" in the advance search section, and the title was searched for springer Link. We searched in Title/Abstract and Keywords and not "ALL" as this often returns a lot of irrelevant articles during the search.

Table 2 documented the search structure separately for each with a specific protocol executed in each database. As shown in Table 2, the Boolean operators were used to include synonyms and alternate spellings while linking enclosed search terms using "OR" and "AND" respectively. The number of results was reduced by utilizing inclusion and exclusion criteria as highlighted below. The reduction was done by selecting articles and conference proceedings published in the English language. The electronic resources as seen in Table 1 were chosen as they cover international scientific sources of high impact-factor (Osadchyi

Table 2 The specific protocol executed in each database

Databases	Protocol
Scopus	(TITLE-ABS-KEY ("machine learning" AND k-12) OR TITLE-ABS-KEY ("artificial Intelligence" AND k-12) OR TITLE-ABS-KEY ("machine learning" AND teaching AND k-12))
Web of Science	TOPIC: ("Machine learning" AND K-12) OR TOPIC: ("artificial Intelligence" AND K-12) OR TOPIC: ("machine learning" AND teaching AND K-12)
IEEE explore	("Abstract":"machine learning" AND "Abstract": k-12 AND "Abstract": schools AND "Abstract": kids) OR ("Abstract":"artificial intelligence" AND "Abstract":k-12) OR ("Abstract":"AI" AND "Abstract":k-12) OR ("Abstract":"artificial intelligence" AND "Abstract":teaching AND "Abstract":k-12) OR ("Abstract":"machine learning" AND "Abstract":teaching AND "Abstract":k-12)
ACM	[[Abstract: "machine learning"] AND [Abstract: k-12] AND [Abstract: schools] AND [Abstract: kids] OR [[Abstract: "artificial intelligence"] AND [Abstract: k-12]] OR [[Abstract: "ai"] AND [Abstract: k-12]] OR [[Abstract: "machine learning"] AND [Abstract: teaching] AND [Abstract: k-12]] OR [[Abstract: "artificial intelligence"] AND [Abstract: teaching] AND [Abstract: k-12]]
Science Direct	(("machine learning" AND "teaching" AND "k-12") AND ("artificial intelligence" AND "teaching" AND "k-12"))
Springer Link	**"machine learning" AND "teaching" AND "k-12" OR "artificial intelligence" AND "teaching" AND "k-12"**

et al., 2020). In addition, Scopus and Web of Science databases provide a strong search functionality (Osadchyi et al., 2020) as well covering many articles. Reference lists of relevant papers were manually searched for additional papers. In an attempt to retrieve more relevant primary studies, the authors also applied a snowball approach and identified 12 additional articles. The authors traced the papers to the database in which it was published and added them to the "Record retrieved" section in Fig. 2.
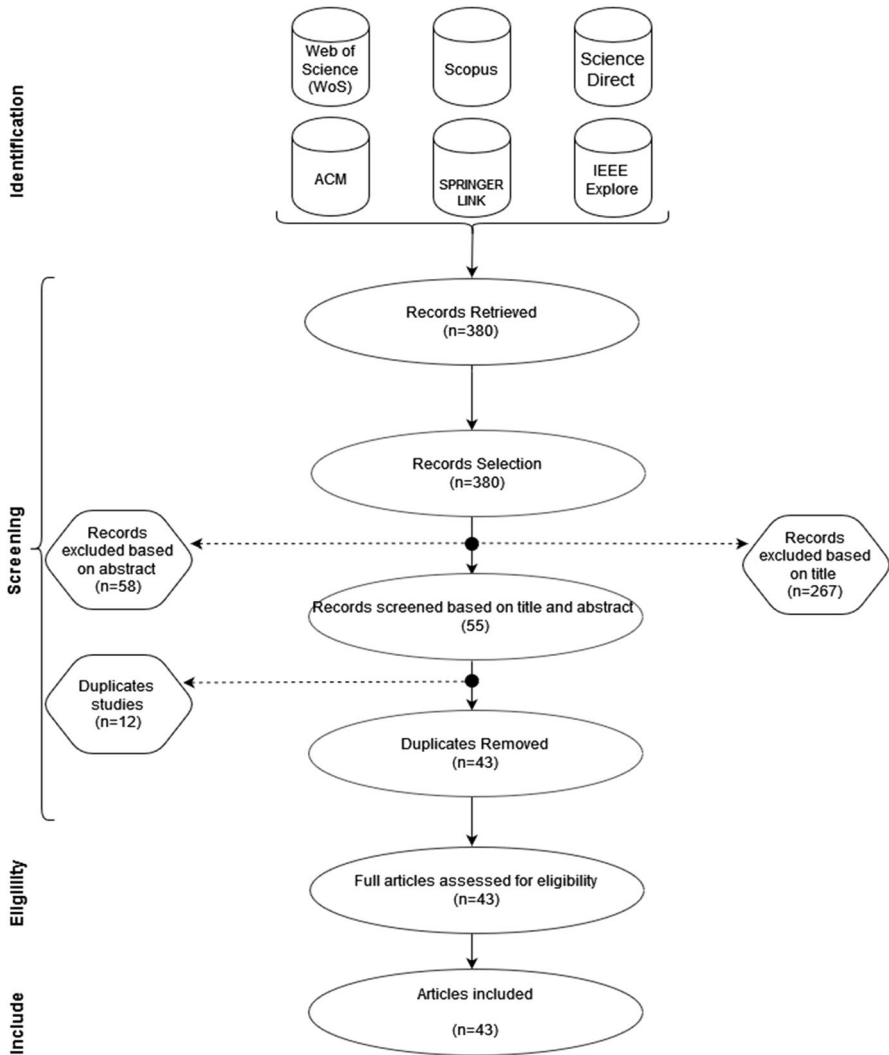

Fig. 2 Overview of processes involved in articles selection

3.2.3 Article selection measures

The search strings in Table 2 were utilized to ascertain that they were pertinent and could provide answers to all the research questions. Shown in Fig. 2 is the process of articles selected in this study which include the criteria for inclusion and exclusion. We read the whole text of the article after which the criteria for inclusion and exclusion were applied to ensure they are relevant for the present research (Table 3).

Inclusion criteria:

- Papers are written in English
- Papers in conferences and journals
- Article that reports teaching and learning of AI/ML in K-12

Exclusion criteria:

- Graduate theses, magazines, letters, notes, and patents
- Duplicate papers
- Articles aimed at the application of machine learning for prediction
- Papers that did not address the teaching and learning of AI/ML in K-12

Articles were included based on the inclusion criteria and excluded based on the exclusion criteria. In total, 43 articles were analyzed in this study haven met the inclusion criteria.

3.2.4 Assessment criteria for study quality

The articles selected after the inclusion–exclusion criteria were evaluated for quality on the quality assessment criteria basis. The quality assessment checklist for this study was adopted from earlier studies (Normadhi et al., 2019; Papamitsiou & Economides, 2014) due to its similarity to the present field of research. As shown in Table 4, a checklist with a 3-dimension Likert scale with dissimilar illustrations for every question, to outline the reviewed literature quality.

Table 3 Number of articles identified at different selection stage

Database	Search results	Analysed articles	Possible relevant articles	Relevant articles
IEEE	13	13	7	7
ACM	198	198	26	10
Science Direct	69	69	13	2
Springer	26	26	10	3
Scopus	20	20	14	12
Web of science	42	42	19	5
Snowballing			12	4
Total				43 (without duplicates)

Table 4 Quality assessment checklist

Item	Assessment Criteria	Description of checklist
1	Does the article evidently define the studies aim?	No, the aim is not described
		Partially, the aim is not clearly described
		Yes, the aim is well described and clear
2	Does paper clearly present a method of teaching or learning of fundamental concepts of Artificial intelligence or machine learning?	No, the teaching of AI or ML is not clearly presented
		Partially, but the teaching methods are not clearly presented and require reading related works
		Yes, it does
3	Is a comprehensive description of the result provided in the studies?	No, the details are not fully described
		Partially, but the details are unclear and must be determined from the references
		Yes, the approaches or strategies can be used with the details as described
4	Does studies present a methodology?	No, no methodology is stated
		Partially, the methodology is present but unclear
		Yes, it does
5	Do the studies receive citations from other researchers?	No
		Partially, 1–5 other articles cite this study
		Yes, more than 5 articles cite this study

3.3 Findings of quality assessment

Table 4 was utilized to gauge the value of the studies included in this study. In question 1, the studies aim of all articles selected was assessed, and all 43 papers evidently detailed their aims while only 18 papers presented the learning aims. The second criterion (QA2) examined whether the studies clearly present the teaching of fundamental concepts of Artificial intelligence or machine learning. Here, 35 papers clearly presented the information concerning the designing, implementation, and evaluation of platforms or activities for teaching fundamental concepts of AI/ML. However, 8 papers included presented an overview of the developed syllabus, curriculum, and guidelines for teaching AI and examined factors that influence students' intention to engage in learning AI. Regarding QA3, we identified 8 papers that reported a researcher-designed curriculum with elaborate descriptions that researchers could adapt or adopt. Fifteen articles described in detail tools existing to teach ML. In addition, 16 papers fully described approaches utilized by earlier research to teach ML while in four papers, the details of teacher training activities were reported described. The 4th criterion is concerned with if the papers obviously show the methodology utilized. The research approach adopted was evidently stated in 28 articles including the methodology utilized (e.g., experimental, design-based, or mixed-method). Five articles mentioned the methodology adopted but clear explanations were lacking. Furthermore, 10 articles basically detailed only their approach but fail to categorically highlight the research methodology that the studies employed. The final criterion focuses on the study's citation in other articles. Google Scholar database was adopted to inspect the number of times the papers were cited (August 2021). Out of the 43 articles reviewed, 15 of the articles received citations more than five times in other papers, 18 received citations rarely (1–5 times) and 8 received no citation at the time of checking while 1 was not found at all on the database. The result from QA5 would at another time differ from the present result as the citations will be updated on the academic database as the papers are cited.

3.3.1 Data extraction and analysis

Based on Kitchenham and Charter's (2007) study, the data were extracted from the papers included with a form. Excel worksheet was designed and finalized afterward having done a thorough revision of the data extracted. The headings sectionalized to retrieve the data are highlighted below:

- Study Objectives
- Resources, materials, and tools
- Learning aims
- Teaching and learning activities
- Issues the studies investigated
- Quantity of articles that refer to the study

An inductive analysis approach was utilized in this study as the concepts are derived from the data. Following Elo and Kyngäs (2008) approach, the process

includes coding and generating categories to offer a means of delineating the phenomenon, to increase understanding and generating knowledge abstraction. The list categories highlighted above are grouped under higher-order headings of curriculum development, technology development, pedagogical development, and teacher training/professional development. In performing the data extraction and analysis process, all the identified articles were considered. Essential data that a paper does not include or evidently stated was allocated "not available" in the matching cell in the table designed for extracting the data.

4 Results

4.1 Overview of included articles

Regarding the publication channels, 39 (90.5%) of the included papers were presented in conference proceedings and only 4 (9.5%) were published in scientific journals.

4.1.1 Summary of papers included according to publication year

As shown in Fig. 3, the summary of included papers according to their year of publication. The papers focusing on K-12 AI/ML increased significantly from 2018 to 2020. Publication on the subject was almost inexistence from 2010 to 2016 and no publication was recorded at all in the year 2017. This trend occurred because the field is still emerging and growing while the high volume of publications produced in 2020 shows increased interest in the research area. The rapid increase of output in 2019 and 2020 suggests more publications in this research area in the coming
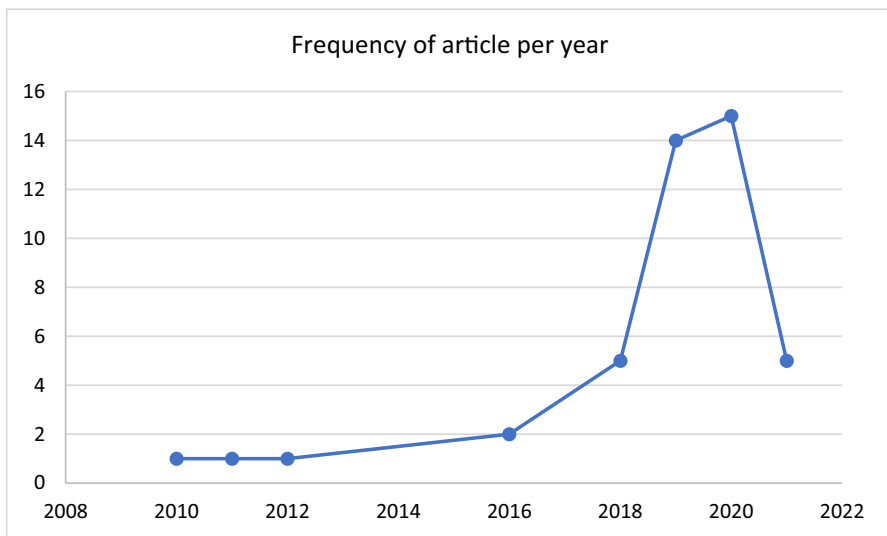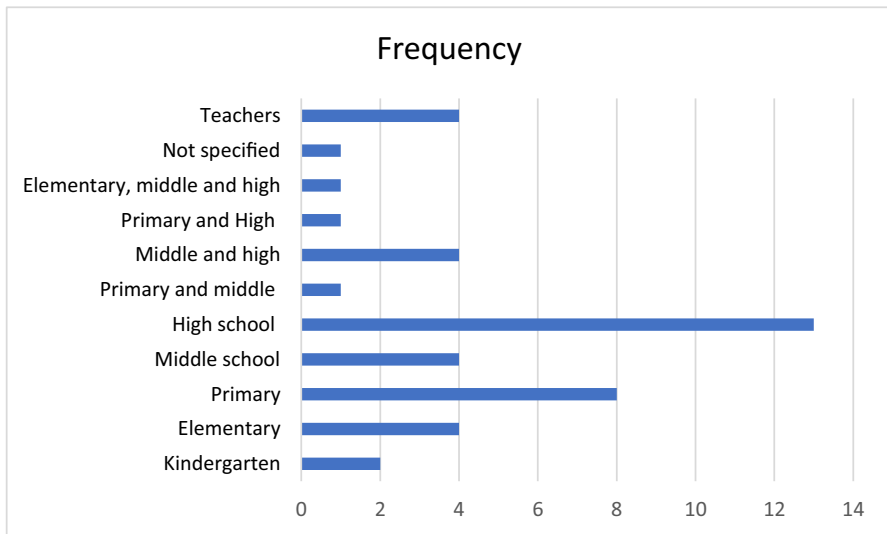

Fig. 3 Articles distributed based on publication year

years due to the activeness of researchers towards the emerging field in recent times as well as the increasing need to democratize AI or ML and involve children. The decline in 2021 could be a result of the time the author conducts the search.

4.1.2 Articles distributed based on educational levels

Of the articles analyzed, the studies were majorly conducted with high schoolers as shown in Fig. 4, n = 13 (Sperling & Lickerman, 2012; Burgsteiner et al., 2016; Mariescu-Istodor and Jormanainen, 2019; Kahn et al., 2018; Wan et al., 2020; Rodríguez-García et al., 2020a, 2020b; Essinger & Rosen, 2011; Ossovski & Brinkmeier, 2019; Evangelista et al., 2018; Lindner et al., 2019; Vachovsky et al., 2016; Estevez et al., 2019). Eight studies focused on primary schoolers (Mariescu-Istodor and Jormanainen, 2019; Lee et al., 2020; Ho & Scadding, 2019; Toivonen, et al., 2020; Chai, et al., 2020; Druga et al., 2019; Tedre, et al., 2020; Hitron, et al., 2018) while only two studies were found that targets Kindergarten (Williams et al., 2019a, 2019b). Four studies focused each on elementary, middle (Sabuncuoglu, 2020; Rodríguez-García et al., 2020a, 2020b, 2021; Sakulkue-akulsuk et al., 2018), middle/high (Opel et al., 2019; Zimmermann-Niefield et al., 2019a, 2019b; Zimmermann-Niefield, et al., 2020) and teachers (Van Brummelen & Lin, 2020; Chiu & Chai, 2020; Kandlhofer et al., 2019; Zhou, et al., 2021) and only one that covers all levels from elementary to high school.

Fig. 4 Distribution of articles based on educational levels

4.1.3 Articles distributed based on educational settings

According to Ainsworth and Eaton (2010), all learning settings, that is formal (e.g., school), non-formal (e.g., museum) and informal (e.g., home), all learning is valuable and contributes to individual`s growth cognitively, emotionally, and socially. As shown in Fig. 5, only one article reported research conducted in an informal educational context, that is, home. More than half of the research were carried out in formal settings (i.e., workshops, course, class sessions) while the remaining studies were conducted in non-formal context (e.g., summer programs, after-school programs). The statistic showing high number of studies carried out in the formal arrangements seem to support the call for inclusion of ML in subject to be adopted in schools.

4.2 RQ1: What is the current focus area of research on ML teaching and learning in K-12?

In the analysis, we identified four categories as derived from the data. The data retrieved from the selected articles are categorized as curriculum development, technology development, pedagogical development and teacher training/professional development as shown in Fig. 6. Hence, we presented our findings according to the stated categories. Curriculum development relates to the content designed to introduce ML and technology development involves the tools developed to teach ML. Pedagogical development focuses on the strategies employed to teach ML while the teacher training and professional development informs of the training possibilities and involvement of teachers in the teaching and learning process of ML.
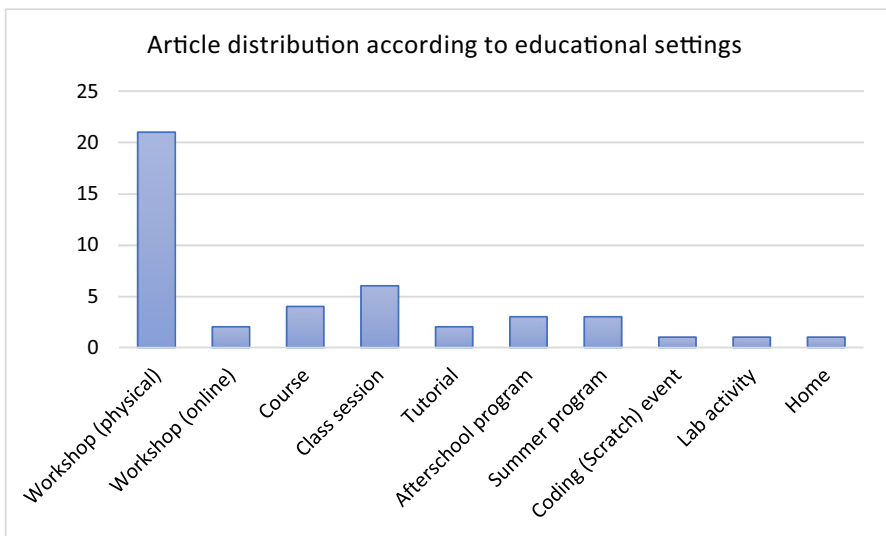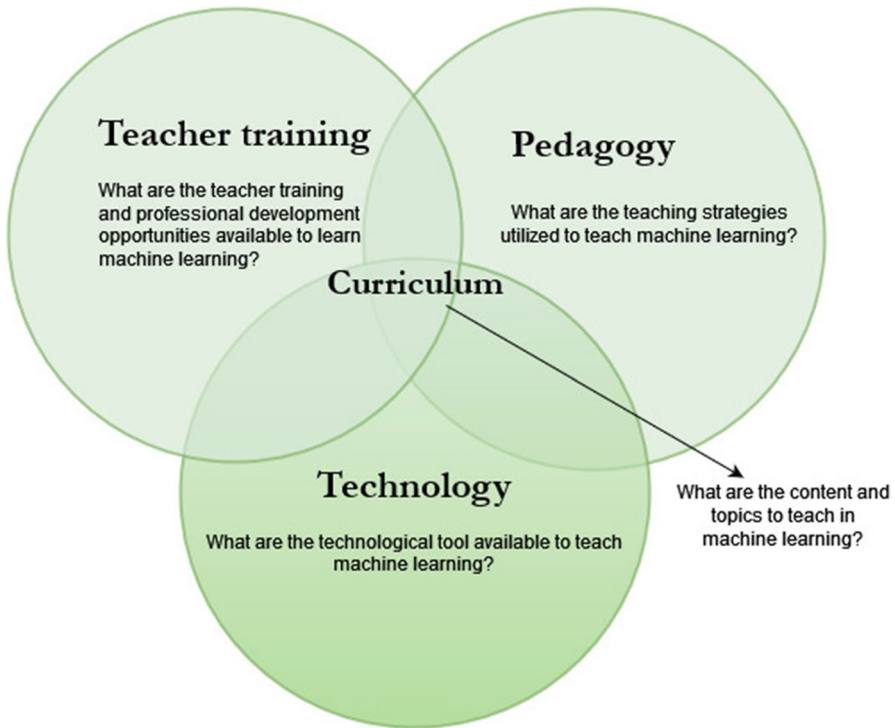

Fig. 5 Distribution of articles based on educational setting

Fig. 6 Focus of the reviewed papers on ML teaching in K-12

4.2.1 Curriculum development

Several curricula were developed by researchers in different regions to introduce ML in schools. Based on our findings, curriculum has been designed to introduce AI concept and promote AI literacy in pre-K and Kindergarten (Williams et al., 2019a, 2019b), elementary (Kim et al., 2021), middle (Sabuncuoglu, 2020) and high (Burgsteiner et al., 2016) school. The objective of the seven studies found focused on how children could grasp the basics of AI. Two studies however stated their learning aim which is that students are expected to learn three AI concepts: knowledge-based systems, supervised machine learning and generative AI (Williams et al., 2019a, 2019b). Most of the articles (4 out of 7) introduced robots as part of the resources to effectively teach AI concepts. These include PopBot which has social robot, made of a smartphone, LEGO blocks, motors, and sensors as components (Williams et al., 2019a, 2019b). Other resources utilized in this category include games, Alexa app or online simulator as well the use of unplugged alternatives. The teaching and learning activities found in this category include the use of Rock, Paper, Scissors games (e.g., Sabuncuoglu, 2020) and role-play as robot games (Heinze et al., 2010). Others include programming exercises, robot construction, discussions, group works, interactive demonstrations (Burgsteiner et al., 2016; Van Brummelen et al., 2020) and riddles and games (Sperling & Lickerman, 2012). The participants were engaged through workshop sessions or an afterschool program.

4.2.2 Technology development

Tools developed for teaching ML considered all grade bands across K-12 levels. The tools identified in the literature are PRIMARYAI, SmileyCluster, AlpacaML, Zhorai, LearningML, and VotestratesML. These tools are mostly web-based platforms designed to assist young children to understand how machines learn. The tools have enabled researchers to introduce ML to students through different fields and interests. For example, Kaspersen et al., (2020) investigated how social studies classrooms can be used as a vehicle to support students' learning and critical reflection about ML. Lee et al. (2020) also focus on integrating AI and life science to teach AI while three other articles explored how to teach youth to make ML models within the space of their athletic interests (Zimmermann-Niefield et al., 2019a, 2019b, 2020). Some of the researchers/developers of the tool specify their learning aim, which is to promote children's understanding of how machines represent knowledge and learning. Wan et al., (2020) specifically used Smiley Cluster to teach k-means clustering to students. Several teaching and learning activities were utilized to introduce the tools to students such as embodied models of gesture (e.g., Zimmermann-Niefield et al., 2020) and the use of mind maps and visualization (e.g., Lin et al., 2020). Other activities include the use of simulation game (e.g., Opel et al., 2019), modelling (Lee et al., 2020), and engaging students in scientific inquiry behaviors such as question asking, and explanation (Wan et al., 2020).

4.2.3 Pedagogical development

It can be deduced from the literature that ML has been taught across K-12 levels from primary to high school. These have also been carried out in various settings such as in laboratory (Essinger & Rosen, 2011), classroom (Chai et al., 2020), summer camp (Narahara & Kobayashi, 2018), workshop (e.g., Druga et al., 2019) and at home (Vartiainen et al., 2020a, 2020b). The objective of the identified studies focuses on how to teach central machine learning concepts with a paper (Vachovsky et al., 2016) concerned with increasing girls' interest in AI. The included articles show that several tools and materials were used in demystifying ML to students. Such tools include Google's Teachable Machine (GTM), robots, scratch, RapidMiner, bounding box, data cards, set of toy cookies and picture cards. Few articles (Essinger & Rosen, 2011; Lindner & Romeike, 2019; Lindner et al., 2019; Ossovski & Brinkmeier, 2019) adopted an unplugged activities approach to convey the central idea of AI to students. Some of the learning aims as deduced from the papers are to teach the concept of facial recognition (Ho & Scadding, 2019), signal processing (Essinger & Rosen, 2011), linear classifier (Ossovski & Brinkmeier, 2019), Decision tree (Lindner & Romeike, 2019; Lindner et al., 2019) and K-mean and ANN (Estevez et al., 2019). To introduce the learning aims to the students, several approaches were utilized. The approaches most adopted include project-based learning which include co-creation of ML-based solutions with the students (e.g., Tedre et al., 2020; Toivonen et al., 2020). Almost all the articles included in this review utilized group work activities to foster students to learn ML basics (e.g., Ossovski & Brinkmeier, 2019; Sakulkueakulsuk et al., 2018). Active-based learning activities were also introduced to the students

through class exercises and facilitated discussions (e.g., Estevez et al., 2019). The use of participatory learning and collaborative methods is also evident in the literature while children co-teach a computer through bodily expressions (Vartiainen et al., 2020a, 2020b). Problem-based learning was also utilized in the study of Essinger and Rosen, (2011) and Evangelista et al. (2018). Lastly, lecture method was utilized to introduce AI concepts among other approaches (e.g., Kim et al., 2021).

4.2.4 Teacher training and professional development

Despite the integral role of teachers in an educational system, our search returns only a few articles focusing on teacher training. This category accounts for the lowest focus among researchers in ML for k-12. Four articles were identified to have specifically focused on teacher training and professional development on AI and ML. These papers include Van Brummelen and Lin (2020) that had a co-design workshop with teachers to identify opportunities to integrate AI education in their course. During the workshop, four different tools were used which includes Machine Learning for Kids, GTM, Google Quick Draw and Google's A to Z of AI cards. Since the co-design was online, Zoom, Slack, and Miro were used to facilitate the sessions. Chiu and Chai (2020) also explore teachers with and without AI teaching experience views on key factors for designing AI curriculum for K-12. Relatedly, Kandlhofer et al. (2019) developed a professional training and certifying system for teachers in AI and robotics. Only an article (Zhou et al., 2021) stood out that focused particularly on ML training for teachers. Zhou et al. (2021) designed a learning platform, SmileyDiscovery, to support low-barrier ML empowered Scientific Discovery without extra ML training for K-12 teachers and learners. The identified papers engaged teachers in the elementary, middle, and high schools with an average of fifteen teachers organized as workshops. The approaches adopted include participatory with the use of real-world applications, gamification, and project-based learning. Three other papers (Lindner and Romeike, 2019; Lindner and Berges, 2020; Sanusi et al., 2022a, 2022b) were found to be focusing on teachers. These papers however gathered the teachers' pre-concepts and perspectives on introducing AI to schools.

4.3 RQ2: What are the issues that should be investigated and future directions of research on ML teaching and learning in K-12?

4.3.1 Issues in teaching machine learning to K-12 leaners

The existing issues in the attempts to democratize and introduce young learners to machine learning through projects and platforms or environments developed were analysed. This study found that not all studies that developed systems reported the limitations in their findings after experiment, tutorial or workshops. The issues were categorised into three dimensions: data, humans and technology as adopted from a study in related field (Normadhi et al., 2019). In this study, the three dimensions categorizes all the shortcomings or limitations of previous research that have been faced while teaching machine learning.

Regarding data issues, previous studies have reported inadequate or paucity of data for confirming the precision, execution, and consistency of machine learning system (Burgsteiner, et al., 2016; Chiu & Chai, 2020; Hitron et al., 2018, 2019; Lin et al., 2020; Van Brummelen & Lin, 2020; Van Brummelen et al., 2020; Wan et al., 2020; Zimmermann-Niefield et al., 2019a, 2019b). According to Normadhi et al. (2019), a small amount of sample size could have enormous influence in signifying a method advances a present system, unravels an issue, or supports a student. Limited duration of experiment was encountered in two previous works (Kahn et al., 2018; Williams et al., 2019a, 2019b). The experiment duration was limited due to the needs of teachers as well the minutes spent by children with each activity or task given to them. Therefore, further work is needed to investigate problem such as time allocation for each session or learning method.

As per the issues connected with human, four problems were identified based on the examined papers: difficulty in learners' understanding ML systems intricacies, overload of assessment and lack of teachers training (Burgsteiner et al., 2016; Chiu & Chai, 2020) and limited learning material/resources (Opel et al., 2019; Sperling & Lickerman, 2012).

This result can be since some participants were unable to conduct their explorations due to technical difficulties (Zimmermann-Niefield et al., 2019a, 2019b), learners were bombarded with numerous assessments or homework involving extensive amounts of time or inadequate skills required by teachers in instructing the learners on machine learning concepts (Ho & Scadding, 2019). Studies such as Opel et al. (2019), Sperling and Lickerman, (2012), Williams et al., (2019a, 2019b), Lin et al. (2020) and Evangelista et al. (2018) had limited resources while recommending learning material in teaching machine learning. For instance, Williams et al., (2019a, 2019b) describe how AI curriculum was designed and assessed its success with 80 Pre-K and Kindergarten children. The study limitations suggest a functional AI curriculum that include more content and can be adapted to other contexts, with learners, novices, pre-service and in-service teachers. Touretzky et al. (2019) recommends conducting research on AI that can be made into an easily available demo, resource, or activity learners and teachers can adapt or adopt. Enabling more materials and resources to know how ML operate and how it will shape their future will increase the student's enthusiasm to explore ML system.

The issues associated with dearth of technology for methods used to develop the learning system were seen in four articles (Essinger & Rosen, 2011; Narahara & Kobayashi, 2018; Zhou et al., 2020), this includes humans and data issue. The suggestions put forward by the identified papers are that the outcome of ML environments be expanded to provide more insight by improving the techniques to improve system performance. The students are also encouraged to design new systems may be through a co-design process involving teachers. Refinement of the system between complexity and usability are another technological limitation (Scheidt & Pulver, 2019; Zimmermann-Niefield et al., 2019a, 2019b). This is proposed to improve the interaction design of the designed systems as well as its user-friendliness on various devices.

5 Discussion

This review elucidates the predominant area of focus in ML for K-12 research in literature. Based on articles included in our review, we discussed our findings under four categories derived from the data. The categories are curriculum development, technology development, pedagogical development and teacher training/professional development. Articles contributing to the four categories are identified and segregated as per their relevance to the categorization. In this section, we elaborated on the review results along as well as its implications with suggestions for further studies.

ML in K-12 education been quite an emerging field currently witnesses exploration among researchers in recent times. We adopted a systematic review approach to understand the present ML studies carried out across K-12 context, recognized the studies that have been conducted, and identified areas thar needs further research. Resulting from our approach, we discovered that most of the reviewed papers of studies leaned towards teaching ML to high schoolers. Literature (Su & Yang, 2022) has shown that limited discussion exists about AI activities appropriate for younger children, however, studies have been carried out to introduce ML at early years of pupils' generation of ML knowledge (Druga et al., 2019; Vartiainen et al., 2020a, 2020b). Nevertheless, no basis could be identified to assume that high schools are the levels that requires ML knowledge, it is essential to conduct research on ML projects suitable for kindergarten, elementary and middle schools. As a result, stakeholders and educational enthusiast could adopt resources valuable to developing students understanding of ML across K-12 levels.

Additional research is necessary for informal contexts to explore activities and interventional studies focusing on ML with students. This is important, especially that learning has been established to be a cumulative process involving connections and support among the variety of learning experiences people encounter in their lives such as at home, schools, and in the community (Dierking et al., 2003). The result of Vartiainen et al. (2020a) study of ML with young children with social support at home suggested that embodied interaction with machine learning systems enhances learning and computational thinking for beginners. Even though some studies were carried out in free choice situations such as Long et al. (2021) that contributes to AI literacy in museum-like settings and online mode (e.g., Druga et al., 2022), we only identified a study (Vartiainen et al., 2020a) that explored ML process with children at home. According to our review, ML studies in K-12 were lacking in informal contexts, as a result, there is a need for more research to understand the contribution of informal-context-based interventions on improving ML.

Third, while the successful integration of topics related to ML into teaching in K-12 requires proper preparation of teachers, teacher training suffers a dearth of research. Lindner et al. (2018) study recognize that teachers had limited understanding and experience to introduce AI, and Sanusi et al., (2022a, 2022b) affirm that teacher training is important to teach ML to students. Hence, ML field requires additional studies to assist teachers in demystifying ML in classrooms. As a result, understanding teachers' readiness to teach ML and their proficiencies could be

relevant. Exploring teachers' readiness to teach ML in classrooms is imperative since teachers' acceptance and disposition could be a pointer to their interest in teaching technology and impact their teaching practices (Nikolopoulou et al., 2021). This assumption is corroborated by a recent study of Ayanwale et al. (2022) who assert that the success of AI education is closely dependent on teachers' readiness and appreciation towards the subject. Based on the above information, teachers cooperation can be ascertained when they are empowerd with the knowledge and approaches to be an advocate of ML in schools.

Fourth, more ML curricula could be designed to integrate ML into science and non-science related subjects. ML could develop learners' knowledge of other subjects, irrespective of the discipline, as well as individual lifestyle (Zimmermann-Niefield et al., 2020; Kaspersen et al., 2021; Lee et al., 2020). While validating effectiveness of ML understanding assist in developing technical and computational skills is required, Kaspersen et al. (2020) emphasized that social studies can be used as a vehicle to support students' learning and critical reflection about ML. In addition, bringing ML into both science and non-science-related subjects would help develop learners for AI-powered world. This could be a solution to the concern of Lin and Van Brummelen (2021) that the dearth of integrated AI curricula across subject domains is a barrier to introducing AI to learners with limited entry to computing-related fields. In addition, an assessment of ML courses could be developed for evaluation of the process when introduced in classrooms. Such of this is a scoring rubric developed by Gresse Von Wangenheim et al. (2022) to evaluate ML learning. Since not all countries use standardized testing, a formative assessment could be designed to represent the students' understanding of ML during the learning process.

Finally, pedagogical and tool development were the most area of focus of the selected studies. Pedagogy development in this paper is concerned with how to teach machine learning in K-12. Since a well-thought-out pedagogy can advance the effectiveness of teaching ML and how students learn, getting to understand the strategies to support learners gain a deeper understanding of the content is essential. Due to the centrality of teaching approaches to simplify ML concepts to children, some papers have emerged focusing on ways to demystify ML at K-12 levels. The pedagogies mostly adopted include group work, project-based learning, activity-based learning, and lecture or instructions. In the reviewed articles, the pedagogical approaches and related theoretical underpinnings were typically not described in detailed manner, which makes it difficult to evaluate how different ML tools were actually integrated into pedagogical practices in diverse settings. Regarding tool development, the role of technology in introducing concepts to learners cannot be overemphasized. This could contribute to why several tools are developed to introduce ML to students. Typically, the visual tools are adopted to teach ML (e.g., Google Teachable Machine or LearningML) that requires no knowledge of programming. Gresse von Wangenheim et al. (2022) reiterated that the tools are mostly available freely online which ensures access and adoption in classrooms depending on internet connections. Curriculum development and especially teacher training/professional development in K-12 ML were currently under-researched and deserved more attention. Given that to enact ML in schools, it must be drawn up in the curricula, it is necessary

to develop more curricula appropriate for different grade levels. Even though some curricula have been developed for several regions and age bands, more should be designed. The need for more curriculum design could be established by a recent UNESCO report that only eleven countries have government-endorsed K-12 AI curricula (Miao & Holmes, 2022).

Even more fundamentally, the dearth of the teacher training program that comprises AI or ML for K-12. With a teacher training program that considers AI/ML, preparing pre-service teachers explicitly trained to communicate AI and ML across K-12 levels is guaranteed (Sanusi et al., 2022a, 2022b). Chiu (2021) also emphasized that teacher AI teaching capacity (e.g., knowledge and pedagogy) is vital to AI education development. A report by UNESCO also highlights teacher training and teacher resource development as essential conditions for supporting AI curricula (Miao & Holmes, 2022). While literature (Mike & Rosenberg-Kima, 2021; Vartiainen et al., 2022) has begun to emerge on pre-service teacher and ML, designing an ML teacher professional program and exploring pre-service teacher's understanding of ML should be a topic for future inquiry. Professional development programs for in-service teachers through hands-on practice workshops and adoption of collaborative design approach, whereby teachers and researchers co-design learning materials and activities could be an effective way to popularize ML in schools.

Practically speaking, as ML continues to be integrated into mainstream education and subject domains across K-12 educational levels, the societal implications including ethics should be emphasized. Ali et al. (2019) *stressed that the* next generation of technologists must be trained to understand the technology ethical and societal impact and not to only see AI as a tool. A study (Skirpan et al., 2018) asserts that infusing ethical dilemmas in the CS curriculum can amplify interest. However, there is currently limited AI/ML ethics curriculum, an exception is Payne (2019) that developed artificial intelligence curriculum centered on ethics for middle schoolers. More ML resources as well as content design considering ethical implications, should be explored, and developed. By so doing, young children will begin to understand the implications of AI technologies they interact with in their everyday life.

Several issues were addressed in the selected studies pertaining to teaching of machine learning as shown in Table 5. Small sample sizes, evaluation of the system, limited learning material/resources, limitations of existing techniques, and duration of the experiment are prevalent inadequacies in teaching machine learning. As a result, this study can benefit future research in cultivating the teaching and learning of machine learning. As a result of the aforementioned challenges in Table 5, our research can inspire developers and practitioners alike to improve the development of ML projects, tools or platforms. The aforementioned works mostly encounter limitations in the appraisal of the methods employed due to number of sample sizes (Burgsteiner et al., 2016; Chiu & Chai, 2020; Hitron et al., 2018, 2019; Lin et al., 2020; Van Brummelen & Lin, 2020; Van Brummelen et al., 2020; Wan et al., 2020; Zimmermann-Niefield et al., 2019a, 2019b). The guiding principle to evaluate students learning of machine learning tools are required for further research. The findings are in line with Marques et al. (2020) particularly in terms of duration of the intervention which can be in form of workshops to short courses and challenge

Table 5 Issues in teaching machine learning

Dimensions	Issues	Bibliography
Data	Limited sample size for system evaluation	Hitron et al., (2019); Van Brummelen and Lin (2020); Burgsteiner et al., (2016); Wan et al., (2020); Chiu and Chai, (2020); Van Brummelen et al., (2020); Zimmermann-Niefield et al., (2019a, 2019b); Lin et al., (2020); Hitron et al., (2018)
	Duration of the experiment was limited	Williams et al., (2019a, 2019b); Kahn et al., (2018)
	Activities completed in groups	Williams et al., (2019a, 2019b)
	Limitations of object in training process	Kahn et al., (2018)
	Data visualization	Zimmermann-Niefield et al., (2019a, 2019b)
	Unanonymized survey	Vachovsky et al., (2016)
Human	Learners' difficulty to understand the intricacies of ML system	Ho and Scadding (2019); Zimmermann-Niefield and et al., (2019a, 2019b)
	Learners feel overloaded with extensive homework	Burgsteiner, et al. (2016)
	Observed a lack of teacher training	Chiu and Chai, (2020)
	Limited learning material/resources	Opel et al., (2019); Sperling and Lickerman (2012); Williams et al., (2019a, 2019b); Lin et al., (2020); Evangelista et al., (2018)
Technology	Limitations of current methods utilized to develop the learning system	Essinger and Rosen (2011); Narahara and Kobayashi, (2018)
	Refining the system between difficulty and accessibility	Zhou and et al., (2020)
		Zimmermann-Niefield et al., (2019a, 2019b); Scheidt and Pulver, (2019)

with evaluating students learning through assessment. More feedback is important to both students and the facilitator to improve the process of teaching and learning.

6 Conclusion and suggestions

This literature review analyzes K-12 ML studies and finds existing gaps and prospective suggestions to conceive and cultivate ML knowledge. Our findings reveal that ML resources were designed and utilized across K-12 settings. The resources were more prevalent in high school than in kindergarten to middle schools. It is also introduced more in the formal school settings than the informal settings. The studies are mainly connected to computing skills which necessitate more research to be conducted on integrating ML into core subjects and domains. While 37 percent of the studies focused on pedagogical development, 35 percent evaluated ML tools. Eighteen (18) percent focused on curricula development and only 9 percent of the selected articles highlight teacher training and professional development. The pedagogical approaches mostly adopted include project-based learning which include the co-creation of ML-based solutions with the students as well as active-based learning activities. As seen in the articles sampled, visual tools (e.g., Google Teachable Machine) (Google, 2020) or LearningML (Rodríguez-García et al., 2020b), or PIC (Tang et al., 2019) are utilized at K-12 schools to demystify ML.

ML for K-12 is an interesting field. This paper aims to aggregate the ML research in K-12 system that have been investigated. Our findings reveal that some effort has been put into research in this field, much more studies is still required. We specifically propose that scholars and K-12 AI/ML practitioners consider the following takeaways when designing ML activities or studies: (a) create more ML activities for kindergarten to middle school, development of teacher training, and education in an informal context; (b) incorporate ML ideas in subject domains other than computing to encourage the integration of ML in schools; (c) develop assessment for ML that can be relevant across levels for students' comparability in ML understanding across learning settings (d) consider societal and ethical implications of ML to better understand students' proficiency of ML; (e) conduct studies focusing on Africa as well as comparative studies exploring different regions to understand how to better introduce ML across climes.

Finally, given that ML has been introduced across K-12 levels, creating a catalog that houses the activities, platforms, and assessments for stakeholders in need of the ML resources would be valuable for the development of the growing field. This could further help to explore the activities or tools in another context with new samples or for teacher training. ML can be integrated into other core subjects and courses across k-12 levels which provides opportunities and shortcomings. Initiating collaborations among practitioners and stakeholders in order to ensure effective ML integration of ML in classrooms. By so doing, ML education could be developed and introduced to students in an effective way and to inspire learners to utilize ML knowledge in other domains.

Recommendations and future directions are proposed based on the findings. Firstly, due to the small sample sizes to evaluate the system mostly

reported in the selected studies, exploring research with larger sample sizes should be done to verify the success of the applied approaches and generalize findings to broader settings (Van Brummelen et al., 2020). Empirical evaluation of studies and project is pertinent even though several appraisal approaches have been utilized (Giannakos et al., 2020) while a solid rubric should be formulated to measure participants' learning gains. Longitudinal pedagogical research should also be considered as it can expand on the understanding of how computational thinking and understanding of machine learning develop in a process of learning (Vartiainen et al., 2020a, 2020b). Furthermore, most studies center on supervised machine learning, particularly on classification tasks, which is regarded as the simplest ML form. Further studies are required to understand if children could comprehend more complex ML processes (Hitron et al., 2019).

Many studies recognize limited learning material or resources, and this suggests that future research should design integrated and AI resources that could be adopted in varying contexts with students, novices, pre-service and in-service teachers. More so, research focuses on the content knowledge as well as instructional design to be adopted for enhancing AI identity and interest (Chiu & Chai, 2020). Lack of teacher training also emerged in the studies, hence, the need for future research to know how to best support teachers to design learning materials and utilize AI technologies into their teaching (Chiu & Chai, 2020). The result of this study would provide valuable data for the teacher training program. Gathering the understandings of teachers and their experiences in co-designing and scaffolding AI in classrooms from various cultural contexts will help generate more useful insight (Chiu & Chai, 2020). In order to democratize access to AI literacy, Wan et al. (2020) recommend co-design workshops with K-12 schoolteachers, Vartiainen et al., (2020a, 2020b) also suggest co-design for children with the support of more experienced peers or adults to build their own machine learning applications. With this, children can gradually develop a deepening understanding of, for instance, different machine learning techniques, data sets, under and over fitting, and testing and improving their systems. Parental involvement in the co-design activities is also encouraged as a future work (Druga et al., 2019; Long et al., 2021). Studies that conduct experiments with relatively long duration are also highly encouraged. It is important because it can lead to discovery rather than mere exposition of the concept. Furthermore, various tools and platforms are introduced to specifically introduce young kids to machine learning and its underlying process. For example, tools such as GTM, ML4Kids, scratch, Popbots and Anycubes. The differences in their applications and the preferences of the learners on various platforms use as well as contents validity is yet to be ascertained. The future studies should hence consider examining the robustness and defect of these tools or projects (Sanusi, 2021a, b). This can inform or suggest learning tools and materials to be selected for future learning activities or specification guides for developing new platforms.

6.1 Limitations

This review has some limitations as common with review papers. This review is firstly limited in search by exploring only six databases. Also, five search terms were utilized which includes "teaching" AND "machine learning" OR "artificial intelligence" AND "K-12" OR "school" in title and keyword. The search terms utilized are frequently used in almost every related reviews. The use of related terms such as "data science" OR "deep learning" may yield more result. The strategy employed in our search may as well influence the presentation of results and be a pointer to limitations regarding generalization. Furthermore, K-12 AI researchers (e.g., Vachovsky et al., 2016) highlighted gender disparity in AI-related education and research. Our study however could not cover this area since not too many reported demographic information. Selecting only English language published articles in journals, and proceedings is also a limitation. It is worthy of note that this work was founded on forty-three (43) paper identified with definite search principles in six databases. Previous studies have concluded based on lesser articles such as Barreto and Benitti, (2012) that reviewed 10 articles and Su and Yang (2022) who concluded based on 17 articles. Though other criteria, strategy and online bibliographic databases may have generated more papers. This research should, hence, be reckoned as an effort to probe into the teaching and learning of ML, rather than a complete overview. We hope that this research will offer valuable suggestions for instructors, practitioners, and researchers in computing and engineering education.

Data availability The authors declare that data supporting the findings of this study are available within the article.

Code availability Not applicable.

Declarations

Competing interests The authors declare that they have no competing interests.

References

Ali, S., Williams, R., Payne B., Park H., & Breazeal C. (2019) *Constructionism, Ethics, and Creativity: Developing Primary and Middle School Artificial Intelligence Education. Presented in the*

International Workshop on Education in Artificial Intelligence K-12 (EDUAI '19). Palo Alto, CA, USA. Proceedings of IJCAI 2019.

Ainsworth, H. L., & Eaton, S. E. (2010). *Formal, Non-Formal and Informal Learning in the Sciences.* Onate Press.

Ayanwale, M. A., Sanusi, I. T., Adelana, P., Aruleba, K., & Oyelere, S. S. (2022). Teachers' readiness and intention to teach artificial intelligence in schools. *Computers and Education: Artificial Intelligence., 1–11,*. https://doi.org/10.1016/j.caeai.2022.100099

Barreto, F., & Benitti, V. (2012). Exploring the educational potential of robotics in schools: A systematic review. *Computers & Education, 58,* 978–988.

Burgsteiner, H., Kandlhofer, M., & Steinbauer, G. (2016). Irobot: Teaching the basics of artificial intelligence in high schools. *Proceedings of the AAAI Conference on Artificial Intelligence, 30*(1).

Chai, C. S., Lin, P. Y., Jong, M. S. Y., Dai, Y., Chiu, T. K., & Huang, B. (2020). Factors Influencing Students' Behavioral Intention to Continue Artificial Intelligence Learning. In *2020 International Symposium on Educational Technology (ISET)* (pp. 147–150). IEEE.

Chiu, T. K. F., & Chai, C. S. (2020). Sustainable Curriculum Planning for Artificial Intelligence Education: A Self-determination Theory Perspective. *Sustainability, 12,* 5568. https://doi.org/10.3390/su12145568

Chiu, T. K. (2021). A holistic approach to the design of artificial intelligence (AI) education for K-12 schools. *TechTrends, 65*(5), 796–807.

Dierking, L. D., Falk, J. H., Rennie, L., Anderson, D., & Ellenbogen, K. (2003). Policy statement of the "informal science education" ad hoc committee. *Journal of Research in Science Teaching, 40*(2), 108–111.

Druga, S. Christoph, F. & Ko, A.J. (2022). *Family as a Third Space for AI Literacies: How do children and parents learn about AI together?.* In Woodstock '18: ACM Symposium on Neural Gaze Detection, June 03–05, 2018, Woodstock, NY. ACM, New York, NY, USA, 17 pages. https://doi.org/10.1145/3491102.3502031.

Druga, S., Vu, S. T., Likhith, E., & Qiu, T. (2019). Inclusive AI literacy for kids around the world. In *Proceedings of FabLearn 2019* (pp. 104–111). https://doi.org/10.1145/3311890.3311904.

Elo, S., & Kyngäs, H. (2008). The Qualitative Content Analysis Process. *Journal of Advanced Nursing, 62*(1), 107–115.

Essinger, S. D., & Rosen, G. L. (2011). An introduction to machine learning for students in secondary education. In *2011 Digital Signal Processing and Signal Processing Education Meeting (DSP/SPE)* (pp. 243–248). IEEE.

Estevez, J., Garate, G., & Graña, M. (2019). Gentle introduction to artificial intelligence for high-school students using scratch. *IEEE Access, 7,* 179027–179036.

Evangelista, I., Blesio, G., & Benatti, E. (2018). Why are we not teaching machine learning at high school? A proposal. In *2018 World Engineering Education Forum-Global Engineering Deans Council (WEEFGEDC)* (pp. 1–6). IEEE.

Georgiopoulos, M., DeMara, R. F., Gonzalez, A. J., Wu, A. S., Mollaghasemi, M., Gelenbe, E., & Alnsour, A. J. (2009). A sustainable model for integrating current topics in machine learning research into the undergraduate curriculum. *IEEE Transactions on Education, 52*(4), 503–512.

Giannakos, M., Voulgari, I., Papavlasopoulou, S., Papamitsiou, Z., & Yannakakis, G. (2020). Games for artificial intelligence and machine learning education: Review and perspectives. *Non-Formal and Informal Science Learning in the ICT Era*, 117–133. https://doi.org/10.1007/978-981-15-6747-6_7.

Google. (2020). *Google Teachable Machine.* https://teachablemachine.withgoogle.com

Gresse Von Wangenheim, C., Alves, N. D. C., Rauber, M. F., Hauck, J. C., & Yeter, I. H. (2022). A Proposal for Performance-based Assessment of the Learning of Machine Learning Concepts and Practices in K-12. *Informatics in Education.* https://doi.org/10.15388/infedu.2022.18

Heinze, C., Haase, J., & Higgins, H. (2010). An action research report from a multi-year approach to teaching artificial intelligence at the K-6 level. *Proceedings of the AAAI Conference on Artificial Intelligence, 24*(3), 1890–1895.

Heys, J. J. (2018). Machine Learning as a Tool to Identify Critical Assignments. *Chemical Engineering Education, 52*(4), 243–250.

Hitron, T., Orlev, Y., Wald, I., Shamir, A., Erel, H., & Zuckerman, O. (2019). 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* (pp. 1–11). CHI 2019.

Hitron, T., Wald, I., Erel, H., & Zuckerman, O. (2018). Introducing children to machine learning concepts through hands-on experience. In *Proceedings of the 17th ACM conference on interaction design and children* (pp. 563–568). ACM.

Ho, J. W., Scadding, M., Kong, S. C., Andone, D., Biswas, G., Hoppe, H. U., & Hsu, T. C. (2019). Classroom activities for teaching artificial intelligence to primary school students. In *Proceedings of international conference on computational thinking education* (pp. 157–159). The Education University of Hong Kong.

Huang, L., & Ma, K. S. (2018). Introducing machine learning to first-year undergraduate engineering students through an authentic and active learning labware. In *2018 IEEE Frontiers in Education Conference (FIE)* (pp. 1–4). IEEE.

Kahn, K. M., Megasari, R., Piantari, E., & Junaeti, E. (2018). AI programming by children using Snap! Block programming in a developing country. In *CEUR Workshop Proceedings* (vol. 2193, pp. 1–14).

Kandlhofer, M., Steinbauer, G., Laßnig, J. P., Baumann, W., Plomer, S., Ballagi, A., & Alfoldi, I. (2019). Enabling the creation of intelligent things: Bringing artificial intelligence and robotics to schools. In *2019 IEEE frontiers in education conference (FIE)* (pp. 1–5). IEEE.

Kaspersen MH, Bilstrup KEK, Van Mechelen M, Hjort A, Bouvin NO, Petersen MG. VotestratesML: a High School Learning Tool for Exploring Machine Learning and its Societal Implications. In: FabLearn Europe /MakeEd 2021-an international conference on computing; 2021. p. 14. 10.1145/3466725.3466728. Design and Making in Education (FabLearn Europe / MakeEd 2021), June 2–3, 2021, St. Gallen, Switzerland. ACM, New York, NY, USA.

Kaspersen, M. H., & Bilstrup, K. E. K. (2020). VotestratesML: Social studies as a vehicle for teaching machine learning. *Constructionism, 2020*, 44.

Kim, S., Jang, Y., Kim, W., Choi, S., Jung, H., Kim, S., & Kim, H. (2021). Why and what to teach: AI curriculum for elementary school. *Proceedings of the AAAI Conference on Artificial Intelligence, 35*(17), 15569–15576.

Kitchenham, B., & Charters, S. (2007). Guidelines for performing systematic literature reviews in software engineering version 2.3. *Engineering, 45*(4ve), 1051. https://doi.org/10.1145/1134285.1134500

Lee, S., Mott, B., Ottenbriet-Leftwich, A., Scribner, A., Taylor, S., Glazewski, K., & Lester, J. (2020). Designing a collaborative game-based learning environment for AI-infused inquiry learning in elementary school classrooms. In *Proceedings of the 2020 ACM conference on innovation and technology in computer science education* (pp. 566–566). ITiCSE '20.

Lin, P., & Van Brummelen, J. (2021). Engaging teachers to Co-design integrated AI curriculum for K-12 classrooms. In *Proceedings of the 2021 CHI conference on human factors in computing systems* (pp. 1–12).

Lindner, A., & Romeike, R. (2019). Teachers' perspectives on artificial intelligence. In *ISSEP 2019 - 12th International conference on informatics in schools: Situation, evaluation and perspectives, local proceedings 2019* (pp. 22–29).

Lindner, A., Seegerer, S., & Romeike, R. (2019). Unplugged activities in the context of AI. In *International conference on informatics in schools: Situation, evolution, and perspectives* (pp. 123–135). Springer.

Lin, P., Van Brummelen, J., Lukin, G., Williams, R., & Breazeal, C. (2020). Zhorai: Designing a Conversational Agent for Children to Explore Machine Learning Concepts. *In Proceedings of the AAAI Conference on Artificial Intelligence, 34*(9), 13381–13388.

Long, D., Blunt, T., & Magerko, B. (2021). Co-designing ai literacy exhibits for informal learning spaces. *Proceedings of the ACM on Human-Computer Interaction, 5*(CSCW2), 1–35.

Mariescu-Istodor, R. and Jormanainen. I. (2019). *Machine Learning for High School Students*. In 19th Koli Calling International Conference on Computing Education Research (Koli Calling '19), November 21–24, 2019, Koli, Finland. ACM, New York, NY, USA, 9 pages. https://doi.org/10.1145/3364510.3364520

Marques, L. S., Gresse von Wangenheim, C., & Hauck, J. C. R. (2020). Teaching Machine Learning in School: A Systematic Mapping of the State of the Art. *Informatics in Education, 19*(2), 283–321.

Miao, F., & Holmes, W. (2022). International forum on AI and education: Ensuring AI as a common good to transform education, 7–8 December; synthesis report.

Michalski, R. S., Carbonell, J. G., & Mitchell, T. M. (2013). *Machine learning: An artificial intelligence approach*. Springer Science & Business Media.

Mike, K., & Rosenberg-Kima, R. B. (2021). Teaching machine learning to computer science preservice teachers: Human vs. machine learning. In *Proceedings of the 52nd ACM technical symposium on computer science education* (pp. 1368–1368). https://doi.org/10.1145/3408877.3439550

Narahara, T., & Kobayashi, Y. (2018). Personalizing homemade bots with plug & play AI for STEAM education. In *SIGGRAPH Asia 2018 Technical Briefs* (pp. 1–4). ACM

Nikolopoulou, K., Gialamas, V., Lavidas, K., & Komis, V. (2021). Teachers' readiness to adopt mobile learning in classrooms: A study in Greece. *Technology, Knowledge and Learning, 26*(1), 53–77.

Normadhi, N. B. A., Shuib, L., Nasir, H. N. M., Bimba, A., Idris, N., & Balakrishnan, V. (2019). Identification of personal traits in adaptive learning environment: Systematic literature review. *Computers & Education, 130*, 168–190.

Opel, S., Schlichtig, M., & Schulte, C. (2019). Developing teaching materials on artificial intelligence by using a simulation game (work in progress). In *Proceedings of the 14th workshop in primary and secondary computing education* (pp. 1–2). ACM

Osadchyi, V., Krasheninnik, I., Spirin, O., Koniukhov, S., & Diuzhikova, T. (2020). Personalized and adaptive ICT enhanced learning: a brief review of research from 2010 to 2019. *CEUR Workshop Proceedings, 2732*, 559–571.

Ossovski, E., & Brinkmeier, M. (2019). Machine learning unplugged-development and evaluation of a workshop about machine learning. In *International Conference on Informatics in Schools: Situation, Evolution, and Perspectives* (pp. 136–146). Springer.

Papamitsiou, Z., & Economides, A. A. (2014). Learning analytics and educational data mining in practice: A systematic literature review of empirical evidence. *Journal of Educational Technology & Society, 17*(4), 49–64.

Payne B.H. (2019). An ethics of artificial intelligence curriculum for middle school students. MIT Media Lab—AI Education. https://aieducation.mit.edu/aiethics.html. Accessed 29 December 2021.

Rattadilok, P. C., & Roadknight, L. Li, (2018). Teaching Students About Machine Learning Through a Gamified Approach. In 2018 IEEE International Conference on Teaching, Assessment, and Learning for Engineering (TALE), pp. 1011–1015.

Rodríguez-García, J. D., Moreno-León, J., Román-González, M., & Robles, G. (2020a). Introducing Artificial Intelligence Fundamentals with LearningML: Artificial Intelligence made easy. In Eighth International Conference on Technological Ecosystems for Enhancing Multiculturality (pp. 18–20).

Rodríguez-García, J. D. D. R., Moreno-León, J., Román-González, M., & Robles, G. (2020b). LearningML: A Tool to Foster Computational Thinking Skills through Practical Artificial Intelligence Projects. *Revista de Educación a Distancia (RED)*, *20*(63).

Rodríguez-García, J. D., Moreno-León, J., Román-González, M., & Robles, G. (2021). Evaluation of an online intervention to teach artificial intelligence with learningML to 10-16-year-old students. In *Proceedings of the 52nd ACM technical symposium on computer science education* (pp. 177–183).

Sabuncuoglu, A. (2020). Designing one year curriculum to teach artificial intelligence for middle school. In *Proceedings of the 2020 ACM Conference on Innovation and Technology in Computer Science Education* (pp. 96–102).

Sakulkueakulsuk, B. S., Witoon, P. Ngarmkajornwiwat, P. Pataranutaporn, W. Surareungchai, P. Pataranutaporn, P. Subsoontorn (2018). "*Kids making AI: Integrating Machine Learning, Gamification, and Social Context in STEM Education*". 2018 IEEE International Conference on Teaching, Assessment, and Learning for Engineering (TALE), 4–7 December 2018, Wollongong, NSW, Australia Page 1005 – 1010.

Sanusi, I. T. (2021a). Teaching machine learning in K-12 education. In *Proceedings of the 17th ACM conference on international computing education research* (pp. 395–397).

Sanusi, I. T. (2021b). Intercontinental evidence on learners' differentials in sense-making of machine learning in schools. In *21st Koli calling international conference on computing education research* (pp. 1–2).

Sanusi, I. T., Oyelere, S. S., & Omidiora, J. O. (2022a). Exploring teachers' preconceptions of teaching machine learning in high school: A preliminary insight from Africa. *Computers and Education Open*. https://doi.org/10.1016/j.caeo.2021.100072

Sanusi, I. T., Olaleye, S. A., Oyelere, S. S., & Dixon, R. A. (2022b). Investigating learners' competencies for artificial intelligence education in an African K-12 setting. *Computers and Education Open, 3*, 100083.

Sanusi, I. T., & Oyelere, S. S. (2020). Pedagogies of Machine Learning in K-12 Context. In *2020 IEEE Frontiers in Education Conference (FIE)* (pp. 1–8). IEEE.

Scheidt, A., & Pulver, T. (2019). Any-Cubes: A Children's Toy for Learning AI: Enhanced Play with Deep Learning and MQTT. *In Proceedings of Mensch Und Computer, 2019*, 893–895.

Skirpan, M., Beard, N., Bhaduri, S., Fiesler, C., & Yeh, T. (2018). Ethics education in context: A case study of novel ethics activities for the CS classroom. In *Proceedings of the 49th ACM Technical Symposium on Computer Science Education* (pp. 940–945).

Sperling, A & Lickerman, D. (2012). *Integrating AI and Machine Learning in Software Engineering Course for High School Students*. ITiCSE'12, July 3–5, 2012, Haifa, Israel.

Su, J., & Yang, W. (2022). Artificial intelligence in early childhood education: A scoping review. Computers and Education. *Artificial Intelligence*, 100049.

Sulmont, E., Patitsas, E., & Cooperstock, J. R. (2019). What is hard about teaching machine learning to nonmajors? Insights from classifying instructors' learning goals. *ACM Transactions on Computing Education (TOCE), 19*(4), 1–16.

Tang, D., Utsumi, Y., & Lao, N. (2019). Pic: A personal image classification webtool for high school students. In *Proceedings of the 2019 IJCAI EduAI Workshop*. IJCAI.

Tedre, M., Vartiainen, H., Kahila, J., Toivonen, T., Jormanainen, I., & Valtonen, T. (2020). Machine learning introduces new perspectives to data agency in K—12 computing education. In *2020 IEEE Frontiers in Education Conference (FIE)* (pp. 1–8). IEEE.

Tedre, M., Toivonen, T., Kahila, J., Vartiainen, H., Valtonen, T., Jormanainen, I., & Pears, A. (2021). Teaching Machine Learning in K–12 Classroom: Pedagogical and Technological Trajectories for Artificial Intelligence Education. *IEEE Access, 9*, 110558–110572.

Toivonen, T., Jormanainen, I., Kahila, J., Tedre, M., Valtonen, T., & Vartiainen, H. (2020). Co-Designing Machine Learning Apps in K–12 With Primary School Children. In 2020 IEEE 20th International Conference on Advanced Learning Technologies (ICALT) (pp. 308–310). IEEE.

Touretzky, D., Gardner-McCune, C., Martin, F., & Seehorn, D. (2019). Envisioning AI for k-12: What should every child know about AI? *In Proceedings of the AAAI Conference on Artificial Intelligence, 33*(1), 9795–9799.

Vachovsky, M. E., Wu, G., Chaturapruek, S., Russakovsky, O., Sommer, R., & Fei-Fei, L. (2016). Toward more gender diversity in CS through an artificial intelligence summer program for high school girls. In *Proceedings of the 47th ACM Technical Symposium on Computing Science Education* (pp. 303–308). SIGCSE '16.

Van Brummelen, J., Heng, T and Tabunshchyk, V. (2020). *Teaching Tech to Talk: K-12 Conversational Artificial Intelligence Literacy Curriculum and Development Tools*. arXiv:2009.05653v1 [cs.CY] 11 Sep 2020

Van Brummelen, J., & Lin, P. (2020). *Engaging teachers to co-design integrated AI curriculum for K-12 classrooms*. arXiv preprint arXiv:2009.11100. arXiv:2009.11100v1

Vartiainen, H., Pellas, L., Kahila, J., Valtonen, T., & Tedre, M. (2022). Pre-service teachers' insights on data agency. *New Media & Society*. https://doi.org/10.1177/14614448221079626

Vartiainen, H., Tedre, M., & Valtonen, T. (2020a). Learning machine learning with very young children: Who is teaching whom? *International Journal of Child-Computer Interaction, 25*, 100182.

Vartiainen, H., Tedre, M., Kahila, J., & Valtonen, T. (2020b). Tensions and trade-offs of participatory learning in the age of machine learning. *Educational Media International, 57*(4), 285–298.

Wan, X., Zhou, X., Ye, Z., Mortensen, C.K. and Bai, Z. (2020). *SmileyCluster: Supporting Accessible Machine Learning in K-12 Scientific Discovery*. Interaction Design and Children June 21–24, 2020 London, United Kingdom Interaction Design and Children (IDC '20), June 21–24, 2020, London, United Kingdom. https://doi.org/10.1145/3392063.3394440.

Williams, R., Park, H. W., & Breazeal, C. (2019a). *A is for artificial intelligence: the impact of artificial intelligence activities on young children's perceptions of robots*. In Proceedings of the 2019a CHI Conference on Human Factors in Computing Systems (pp. 1–11).

Williams, R., Park, H. W., Oh, L., & Breazeal, C. (2019b). Popbots: Designing an artificial intelligence curriculum for early childhood education. *In Proceedings of the AAAI Conference on Artificial Intelligence, 33*(1), 9729–9736.

Zhou, X., Van Brummelen, J., & Lin, P. (2020). *Designing AI learning experiences for K-12: Emerging works, future opportunities and a design framework*. arXiv preprint arXiv:2009.10228. arXiv:2009.10228.

Zhou, X., Tang, J., Daley, M., Ahmad, S., & Bai, Z. (2021). "Now, I want to teach it for real!": Introducing machine learning as a scientific discovery tool for K-12 teachers. In *International conference on artificial intelligence in education* (pp. 486–499). Springer.

Zimmermann-Niefield, A., Polson, S., Moreno, C., & Shapiro, R. B. (2020). Youth making machine learning models for gesture-controlled interactive media. In *Proceedings of the interaction design and children conference* (pp. 63–74). IDC'20

Zimmermann-Niefield, A., Shapiro, R. B., & Kane, S. (2019a). Sports and machine learning: How young people can use data from their own bodies to learn about machine learning. *XRDS: Crossroads, the ACM Magazine for students, 25*(4), 44–49.

Zimmermann-Niefield, A., Turner, M., Murphy, B., Kane, S. K., & Shapiro, R. B. (2019b). Youth learning machine learning through building models of athletic moves. In *Proceedings of the 18th ACM international conference on interaction design and children* (pp. 121–132). IDC '19.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Authors and Affiliations

Ismaila Temitayo Sanusi[1] ⬤ · **Solomon Sunday Oyelere[2]** ⬤ ·
Henriikka Vartiainen[3] ⬤ · **Jarkko Suhonen[1]** ⬤ · **Markku Tukiainen[1]** ⬤

Solomon Sunday Oyelere
solomon.oyelere@ltu.se

Henriikka Vartiainen
henriikka.vartiainen@uef.fi

Jarkko Suhonen
jarkko.suhonen@uef.fi

Markku Tukiainen
markku.tukiainen@uef.fi

[1]　School of Computing, University of Eastern Finland, P.O.Box 111, 80101 Joensuu, Finland

[2]　Department of Computer Science, Electrical and Space Engineering, Luleå University of Technology, 931 87 Skellefteå, Sweden

[3]　School of Applied Educational Science and Teacher Education, PO Box 111, 80101 Joensuu, Finland