

Equity and Bias in AI Educational Tools: A Critical Examination of Algorithmic Decision-Making in Classrooms

Shumaila Farheen¹, Azhar Abbas Cheema², Rai Samee Ullah³ & Dr. Marium Minhas Bandiali⁴

¹MPhil, Department of Education, Email: Shumaila14786@gmail.com

²MS Management Sciences, University of Sargodha, Email: azharprsp@gmail.com

³Department of Computer Science and Software Engineering, Grand Asian University Sialkot, Email: sameetatla@gmail.com

⁴PhD Public Policy, Deputy Director (T/C), Directorate of Sindh water sector & barrages improvement, GoS, Email: marium_minhas@hotmail.com

ARTICLE INFO

Article History:

Received:	May	12, 2025
Revised:	June	15, 2025
Accepted:	June	23, 2025
Available Online:	July	01, 2025

Keywords:

Artificial Intelligence, Algorithmic biasness, Educational Equity, AI in Education, Teachers' Perceptions, AI Experience, Quantitative Research, Educational Technology, Decision-Making, Pakistan.

Corresponding Author:

Shumaila Farheen

Email:

shumaila14786@gmail.com



ABSTRACT

The integration of artificial intelligence (AI) into education raises significant concerns about algorithmic bias and its impact on equity. This study examines the linkage between teachers' perceptions of AI bias and their concerns for fairness in the classroom while also assessing the influence of teaching level and AI familiarity. Through a cross-sectional survey of 270 educators in Punjab, Pakistan, we utilized quantitative methods, including correlation and regression analysis, to test our hypotheses. Our findings reveal a strong positive association between perceived algorithmic bias and threats to educational equity. We found that school teachers were significantly more aware of AI bias than university faculty. Furthermore, greater teacher familiarity with AI tools correlated with a more nuanced understanding of their potential biases. These results underscore the urgent need to address the "black box" nature of educational AI. The study provides empirical evidence that, without careful regulation and comprehensive teacher training, AI-driven tools risk perpetuating and even exacerbating existing educational inequalities, aligning with global concerns about the ethical deployment of AI in sensitive sectors.

Introduction

Artificial intelligence (AI) is one of the most popular directions in the sphere of education that introduces both challenges and opportunities to modern teaching (Pedro et al., 2019). Artificially intelligent applications (AI in learning), such as intelligent tutoring systems, automated grading programs, and predictive analytics, have transformed the manner in which teaching and learning

are conducted and assessed and even administrative choices within education. Implementation of such systems may lead to efficiency gain in learning, as the system will permit the establishment of a streamlined learning environment, provide instant feedback, and deliver the contents on demand based on the competencies and preferences of individual learners. However, as the use of AI as a tool for facilitating the most basic school operations grows more and more popular, the transparency, fairness, and responsibility of algorithms also become of growing importance. Much of the concern of teachers is that such cloud-based systems do not ensure an equal learning process, and it may result in the appearance of new forms of social inequality since the decisions gained with the help of algorithms may still be influenced by some intangible bias.

The rationale in this case is that the judgments passed on the AI systems are not always transparent, and thereby determining whether such interventions can be fair or not is hard. This is especially the case with asymmetries like race, sex, socio-economic characteristics, and language proficiency, to mention but a few. One of the dangers that is supposed to heavily influence the switch to the implementation of technologies powered by AI by educational organizations is algorithmic discrimination. This is one of the particular concerns in a multicultural learning environment where cultural and demographic variations are crucial elements that determine the learning process. Its members, in their turn, are growing even more assertive to imply that even the use of AI tools could not be sufficient to base it on their sufficiency of operations, and social and ethical implications would have to be sorted out (Holstein & Doroudi, 2021). The biasness of the algorithms has its origins in the materials on which these machines have been trained or the assumptions incorporated into the algorithms themselves. The application of AI algorithms in educational settings can cause discrimination by reinforcing a historical biasness and inequality, which ultimately will reproduce the markers of inequality in the past.

To take a simple example, consider predictive analytics algorithms best suited to predicting the risk of a student; in these circumstances the algorithm can wrongly classify students who fall under minority or marginalized demographics as at risk not because of the situation in reality, but because of biased data covariates. The fact that guidelines and policies applied through AI can contain prejudices and a shortage of equity initiatives can be far more intrusive to teachers and policy leaders than the notion of having to be objective and technologically more proficient. Further, the fact that users cannot refute or contradict an infallible outcome or one that is highly probable to be incorrect because of the transparent way of decision-making that AI is often presented as in the literature can be explained by the black box problem. The matter is especially critical in learning environments, where fairness is one of the main principles and moral priorities. To tackle the issue of algorithmic biasness, the professionals need to consider thoroughly the sources of used data, structures of models, and strategy of implementing the AI-based educational resources. Moreover, policy frameworks and ethical principles that are necessary in guaranteeing fairness, transparency, and inclusiveness of all learners in the application of AI in education are advancing (Stephenson & Harvey, 2022).

When entering the sphere of implementation of AI in the practice of assigning progression, evaluation, and prescribing the same interventions without an adequate focus on the interaction of people, one risks paying a price that is quite high: simple reduction in the particular educational process itself and its avouchment to the automated process instead of a professional decision-making procedure in the form of the teacher. Also, one can mention that such risks as the situation of being powerless to the students when the algorithm-related decisions are applied to them may be encountered when the students have not received any information about the system, and they are not provided with the chance to disagree with the conclusion and/or to appeal the decisions that impact their academic achievements. A problem of data privacy and surveillance issues may also

be included in the list as the AI systems continually collect, process, and act on the fundamental and behavioral data of students. The given dilemmas reveal that it is essential to substantially control the AI in education because these technologies will be deployed to facilitate the mission of the democratic society and needs of social justice. The ethics in support of fairness, inclusion, and the well-being of the student need emphasis as the introduction of AI in education is rapid. Without those efforts, it is most likely that one will only continue the inequalities that exist and fail to trust an institution of higher learning (Rasool et al., 2025).

Understanding AI Integration in Educational Settings

One of the most significant technological innovations in the classrooms of the contemporary period is the use of artificial intelligence (AI) in the classroom. The AI technologies have already been integrated into various educational processes, including individual tutoring systems, automatic graders, and student retention predictive analytics. As educational facilities switch towards using AI, there are a lot of problems entailed in the AI development in the education context: it may be fairness, inclusivity, transparency, or unintended consequences. In that regard, in order to evaluate the applicability of AI tools in education, one requires knowledge of how the tools operate in an education setting to adopt the use of AI tools in education in a responsible manner. Furthermore, the publication of regulatory standards and the principles of ethics in AI that apply to an educational setting is another significant process to uphold the rights of the students and the liability of the institution. Communication between educators, policymakers, and the AI developers should be maintained all the time to ensure that there is no erosion of the teaching essence of education with the use of AI as opposed to enhancing it (Khosravi et al., 2022).

Applications of AI in the learning environment not only end in the instructional application, which is in widespread application, but also in the administrative level and the policy level. Admission of students, school attendance monitoring, academic advice, and course recommendation, among others, are some of the school practices that schools around the world are adopting AI to support. Due to the power of the learning management system, the learning management system can monitor the online activity pattern of students; hence, by the time they become at risk of dropping out, they are identified and interventions suggested. Though this type of application has the tremendous potential to enhance educational services, it is linked with the possibility of expanding the status quo inequalities in case algorithms operate the prejudiced databases. Moreover, AI-based systems are not always transparent at all, and the ones studying and teaching there do not see how decisions can be reached and why some solutions have the logic they hold. Opacity and the out-of-sight, out-of-mind effect of the mundane operation of algorithmic action complicate the culpability of such systems, combined with the fact that in the suggested AI-based recommendations, the reasoning chain is often not available. Researchers have always indicated that the explain ability and transparency of the AI applications to be used within education are the main issues that contribute to sustaining the credibility, goodwill, and justice of the decision-making processes. Without this protection, AI may become an absolute power in the classrooms, killing the process of human regulation and professional experience. Institutions of learning ought to therefore come up with clear ethical norms and evaluation systems to monitor the impacts of the analyzed tools of AI on students and their learning activities and the institutions practices (Holstein & Doroudi, 2022).

When the AI tools are trained on the culturally homogenous material, the systems will fail to identify those pattern-rich divers. In addition, such an overwhelming amount of quantitative values of educational accomplishments, i.e., the test scores or engagement rates, of course, neglects the qualitative spheres of education, like creativity, critical thinking, and emotional health, which are

equally essential aspects of the educational process, as well as the quantitative one. As the educational systems continue their trend toward adapting to the AI-based individualistic paradigm of learning, one should ask the questions, who will benefit in the implementation of the technologies, and who can be left behind? This requirement proves the importance of enabling all-inclusive representative data practices and models and creating accountable and ethical AI in learning. What is more, it ought to be mentioned that by engaging educators in the development and evaluation of AI tools, it is possible to assume that the technological innovations will not conflict with the pedagogical values and will not be associated with the achievement of the targets related to equality (Holstein & Doroudi, 2021).

As per research studies, more teachers are of the opinion that AI may enhance learning, but they also feel that AI is inaccurate and inequitable and that it also harbors ethical problems. Educators are concerned with the prospect of AI-based judgment undermining the academic success and emotional well-being of their students and fear losing control over the decision-making procedure. That is why, in order to effectively implement AI in education, it is not only necessary to prepare a technology foundation but also professional development programs that would facilitate appropriate training and skills/knowledge provision needed to employ AI-powered classrooms successfully. Further on, these issues are critically important to establish confidence of teachers in AI systems and make them supportive ones instead of substitutes for human judgment. They also require continuous research and consultation in order to evaluate the long-term impact of the integration of AI into the teaching practices and the equity of students (Rahmanipur & Mohammadi, 2025).

Algorithmic Biasness in AI Educational Tools

Algorithmic bias can be defined as the production of undesirable and unfair results by AI applications in a typically systemic way. This is partially because these applications rely on social injustices that have already been established and that they have learned about during training. There are two ways in which this bias can be incorporated: first, it can be present when students' learning outcomes are interpreted unfairly due to differences in socioeconomic status, gender, race, or language ability, and students who belong to a particular group are routinely identified by these devices. A good example of the application of AI-based educational tools, like predictive analytics and automated grading systems, often fill in their historical data, which can perpetually re-enforce pre-existing biases within machine-learning algorithms. What makes this issue more serious is that an extremely limited number of persons can review, understand, and contest the logic of making a decision AI-generated, such as educators and administrators (Dwivedi et al., 2023). To look at what experiences, views, and groups are prioritized or left behind by AI systems, the researchers note a pressing demand to develop a critical analysis of such systems. In addition, it is necessary to create awareness regarding algorithmic bias among teachers and students in order to eventually make AI use in education responsible and ethical. In the absence of such efforts, such technologies would lead to solidifying the problem of inequality and fail to achieve the primary idea of fairness and inclusivity of educational settings (Binns, 2022). It is determined that data sets and learning algorithms that form the foundation of AI systems are the source of algorithmic bias.

Any prejudice, including achievement gaps, resource disparity, cultural bias, and other inequities, may be documented in educational data that is commonly fed into AI algorithms. Prejudice permeates the system when you use such biased data to pretrain the AI tools without making any conscious attempts to improve the situation. More importantly, because predictive algorithms would deny kids the opportunity to enroll in advanced learning programs, the risk-habit classification of low-income students is likely to be erroneous. Additionally, AI-based algorithms

frequently overvalue quantitative metrics like test scores and attendance grades while undervaluing qualitative traits like creativity, leadership, resilience, and critical thinking, which are frequently overlooked by more conventional grading schemes. The researchers have underlined that AI bias could be concealed behind the appearance of objectivity, presenting unfair outcomes with elements of unsocial notions masquerading as data-driven realities. Educational equity may be seriously jeopardized by inactive equity checks and infrequent privacy audits of AI-driven systems, as the biased algorithm may unintentionally replicate existing disparities. Therefore, it is crucial to employ fairness assessment tools and incorporate more varied and representative datasets into AI development in order to reduce potential harm and make inclusive teaching practices viable (Hintz, 2024).

The algorithm's difficulty is the second barrier to comprehending and resolving bias in systems that use AI for learning; the more people grasp this, the more challenging it is to confront and resolve the issue. For example, a computer program used to evaluate essays is unable to identify the writing styles or argument structures that students from diverse cultural backgrounds frequently employ, which reduces the likelihood that they would receive good ratings. Moreover, the ability of systemic discrimination is provoked by the presence of the network of multiple AI systems used in schools, e.g., behavioral-tracking, evaluation, and attendance-tracking systems. Researchers define that explainable AI is needed in schools in order to make sure that decision-making processes are transparent and available to both the teachers and the students. Explainability gives confidence in AI-based educational decision-making and contributes to realizing justice (Holstein & Doroudi, 2019). Therefore, the problem of educational AI tools is one of a highly institutionalized nature rather than one that is only technological or even socially oriented. More academics who support the notion of equitable participation in the creation of AI systems agree that more students, teachers from diverse backgrounds, and other community people should be involved. In addition to encouraging the development of more egalitarian, inclusive, and open spaces, it is hoped that these pro-inclusive initiatives would help to recognize, denounce, and reverse long-standing biases (Williamson & Eynon, 2020).

Bias and Inequity in AI Decision-Making

AI-based learning environments have a strong impact on the academic and emotional maturity of students affected by discrimination. Eventually, the AI technologies will unwillingly steer inequality against given populations, thus continuing the problems of existing educational inequities (Bulathwela et al., 2024). Such disadvantaged populations, such as the members of ethnic minority groups, people with low incomes, and students with disabilities, are more likely to undergo disproportionate effects of AI-made decisions in comparison with those who are not severely marginalized. One of the examples of how this may happen is by the use of predictive analytics programs. It is based on these programs that such students are labeled as at risk due to prejudicial information that has been provided in the past and may lead to unnecessary intervention or even exclude them from most of the learning environments at their disposal. Prejudice in AI is holding these schools in a snag and making itself the normal thing, so ensuring they go into a loop of failure in the case of pessimism, something they must purposefully get out of. The scientists emphasize the significance of the fact that, unless the tendency of using them in this sphere is reversed, artificial intelligence systems are to be mere reductions in lieu of the unjust decisions in the name of objectivity, stiffening the fabric of inappropriateness instead of enhancing it (Binns & Kirkham, 2021).

Bias in the proposal of AI systems is not an exception since these systems are indeed developed without sufficient knowledge of the social, cultural, and contextual diversity of pupils. The

implementation of AI tools is also likely to rely on data that do not take into account the specific circumstances of learners, their background, and their needs in knowledge. As such, the suggestions of AI will not suit the needs of the pupils that do not have the same socioeconomic backgrounds, methods of learning, and cultural backgrounds as the training data. This difference in the notion portrayed by AI systems and their reality may relegate students and misclassify them out of context, i.e., they may be treated undeservedly and classified differently due to the difference between what AI systems show and the reality of the situation. In the example of the AI-enabled attendance tracking systems, the impoverished children will be penalized in case they have to take care of their naughty children or commute long distances. The socioeconomic facts that define the way students participate and particular engagement are not considered with the ways. The developers, legislators, and teachers must understand that developing context-aware algorithms that can appreciate the diversities of every student is important to attaining equity in AI. Additionally, the scholars back the idea that the regulatory structures should be developed to include moral and ethical factors along with technical performance aspects in the design of the AI systems. It is also important to believe in the constant contribution in terms of feedback about various groups of students in order to make sure that, by means of these technologies, inclusive and socially responsible approaches are offered (Whittaker et al., 2019).

It is another critical implication of bias in AI decision-making to diminish trusting relationships among students, teachers, and institutional leaders. Inability to perceive the AI systems as being fair or without any discrimination can demoralize the students in both the traditional schools and the online schools. The prejudice in the results carried out by an AI can result in a feeling of exclusion and alienation, especially when students believe that algorithms make random or biased decisions concerning their skills and capabilities. To educators, biased AI tools negatively affect their professional autonomy because they reduce their competency to safeguard algebraic offerings with their skillfulness and knowledge of students. This compromises the autonomy of teachers and instead requires them to defer to secretive, unanswerable AI. The researchers always say that trust is a necessary component of learning and teaching, and in order to achieve trust, AI systems have to be transparent, fair, and explainable. Devoid of such properties, AI can turn out to be a disruptive and dividing phenomenon in the educational setting instead of a helpful learning aid. Besides, regaining lost trust may be a process that takes a long time, and it is important that these issues are handled before they occur. Participatory design should be used, and educators should be engaged in developing AI systems to match the technological advances with the pedagogical interests and be student-friendly (Holstein et al., 2019).

Research Objectives

1. To critically examine the presence and nature of algorithmic bias in AI-driven educational tools and its impact on decision-making processes within classrooms.
2. To explore the implications of AI-based inequitable decision-making on students' academic opportunities, emotional well-being, and educational experiences.
3. To propose practical, equity-focused strategies for minimizing bias in AI educational systems through inclusive design, data fairness audits, and transparent policy frameworks.

Problem Statement

The use of artificial intelligence (AI) tools in a classroom setting is becoming more frequent as it transforms the understanding of decision-making in the learning environment, giving learners a chance at personalized settings and granting administrative advantages. Nevertheless, such fast growth has caused a lot of concern due to the danger of programmed inequality and discrimination

in AI-based learning. Historical data and non-transparent software algorithms form the basis of AI systems, which are able to recreate the current inequalities in terms of social, cultural, and academic practices, which could influence the underrepresented and marginalized student groups the most. Even though the level of AI in classrooms appears to be increasing, no critical consideration has been given to how these new methods of teaching can affect all the academic experiences, the way they can open the choices of resources, and student achievement. This gap should be filled, and AI systems should serve as vehicles of equity and inclusion in education rather than engines of inequities.

Significance of the Study

This study is valuable since it will be used in the emerging discussion regarding AI ethics as well as equity and justice in edtech. By presenting the problem of algorithmic bias in classroom decision-making supported by AI as the issue of concern in educational policy and practice, the study responds to a timely demand in the sphere of education. It is also an indication that educational establishments and technology companies must be required to make it certain that further development of AI does not lead to repeating or enhancing any present disparities. When creating guidelines, policies, and ethical frameworks that may contribute to the transparent, fair, and culturally sensitive use of AI in educational environments, the findings of the present study may be employed. Also, the research cites the support of education for the students and educators under the influence of AI-deployed systems and highlights the importance of a participatory process setting of AI design and fairness audit. This report can be of use both to the policymaking writers and developers of technology as well as to teachers and academic establishments anywhere in the world since it provides knowledge on how educational justice can be realized in the future by utilizing such AI tools. Its recommendations can be helpful in transforming AI technologies into instruments of inclusiveness, fairness, and ethical responsibility in education.

Literature Review

Through this fast deployment of artificial intelligence (AI) in education, new lines have been opened to personalized learning, data-guided instruction, and administrative effectiveness. Nevertheless, along with these short-term advantages, a number of ethical and social questions have been raised when using AI-based decision-making, especially in terms of algorithm bias and its role in educational inequity. Researchers claim that a common pattern is that AI systems receive the biases that are present in the data of the past and that are reproduced in the AI systems, meaning that the outcomes are disproportionately in favor of or against groups of students, most often based on race, gender, socio-economic status, or cultural background (Holstein et al., 2019). The biases are generally underpinned by the shortcomings of training data, which tend to reflect the past social inequalities and do not reflect a diverse diversity of present-day classrooms. With the increasing role that AI has played in establishing the grades, grading of students, and also distribution of resources in an educational institution, the issue of its ability to recreate the system inequities within an educational system has become the topic of growing academic interest (Binns, 2024).

The implications of unfair decision-making on the basis of AI extend further past the academic performance and reach the emotional stability and overall experiences of the students.

There is evidence that active defiance against the system of AI is a major part of demoralizing students, negatively affecting their self-esteem in academics, and making them feel disadvantaged in education. In the past, predictive analytics systems also exhibited the kernel of marking the students with underprivileged backgrounds as being high-risk due to the presence of the biased

historical data, which has created a situation of the over-intervention or over-restriction of the students in question, limiting access to the learning opportunities (Swist & Gulson, 2023). Scholars have proposals of participatory AI design in which communities, educators, and students can be involved in developing and testing AI tools. In addition to this, as an optional measure, the fairness audits and bias detector mechanisms should be proposed, which will allow getting rid of such disparity as that identified in the results generated by AI before the latter affects the educational processes of the students. It is also imperative to devise transparent policy systems and rules to guide the ethical application of AI in learning and education such that it can be fitted within the guideline of inclusiveness, justice, and humankind learning guidance (Veale et al., 2018). Whether AI is relevant in the classroom environment, researchers have held that data-driven systems, in the latter case, AI-enabled applications, can be applied to detect the at-risk pupil even before it is very late and thwart them through offering them specific assistance in real-time predicated on real-time analysis (Chinta et al., 2024). However, even though, based on assumptions, the perceived learning benefits of such tools should be quite costly, a growing body of evidence is suggesting that the complex of ethical and equity-based apprehensions is associated with its use, including the problem of the bias and equity of the algorithms in particular. Research has demonstrated that the biases that exist in society in the forms of current data sets are likely to be reflected and augmented by AI tools that were built based on the same data sets (Ntoutsis et al., 2020). As an illustration, the predictive models applied in grading and monitoring of the students have shown different performances depending on the socio-economic status, race, gender, and cultural background (Binns, 2018). Most AI-based decision-making systems are considered as black boxes, having a minimal explanation of how certain results or categories are arrived at.

This is non-transparency that cannot allow educators, students, as well as parents to comprehend, query, or oppose decisions made by these systems (Barnes & Hutson, 2024), stressed that unless explainable and interpretable, AI will pose a danger to the trust in educational institutions and will increase the disparities that already exist. In a bid to address the dilemma, Williamson and (Eynon, 2023), demanded a more comprehensive study of the use of AI in education that should consider the input of the social justice system, ethics, and educational psychology. The fact that one should review the implications and consequences of the AI-based decision-making critically means further research will make it possible to introduce even more just and accommodative strategies regarding AI development and classroom application. Such works of background literature will be discussed in the next literature review, yet they will also consider the gaps that still exist; hence, they will be able to tackle the issue of inequity in regard to the application of AI tools along with providing some recommendations on how to address the problem of bias in learning technology.

Algorithmic Bias and Equity Challenges in AI-Driven Educational Systems

Such AI-based technologies as intelligent tutoring systems, predictive analytics, and automated grading software have allowed educators to enhance learning by supporting the individualization of educational experiences, automation of clerical procedures, and well-balanced efficiency within classrooms and university settings (Pedro et al., 2019). The ability to provide individualized content to students according to their requirements, track academic performance of students in real-time, and automate intricate data analyses that have traditionally been handled by schooling personnel are all hailed as the key features of these systems. Nevertheless, rather unstructured and unregulated employment of AI in the educational sector has also been accompanied by ethical issues related more to issues of justice, openness, and inclusiveness (Khosravi, 2025). Scientists believe that AI technologies are more focused on quantitative academic outcomes, including test-score performance, attendance history, and problem-solving rates, whereas AI-based technologies

usually do not take into account qualitative indicators, including critical-thinking skills, emotional intelligence, and a student with an innovative business (Doroudi, 2024).

Consequently, the educational contexts relying more on data-driven models can easily ignore all socio-economic, cultural, and psychological conditions dictating the experiences and academic outcomes of the learners and thus, promote inequities in the AI-mediated decision-making process (Eynon, 2022). By relying on these biased sources of data as inputs in predictive models, including early warning systems or performance risk assessment, the latter may reproduce (or even compound) the prior disparity within the learning institutions (Shen et al., 2021), such extensively described instance is that of predictive analytics tools, which overrate minority or marginalized students as underperformers, despite their potential application, simply because of data covariates biased on history and the inequities of systems that exist (Williamson et al., 2023).

These consequences limit access to academic opportunities for disadvantaged groups themselves but, more importantly, perpetuate negative stereotypes, which continue to strengthen existing systemic inequities within the educational context (Yeung & Harkens, 2023). For instance, students whose problem-solving techniques or textual communication do not conform to dominant or culturally mainstream norms may receive lower scores that are manipulated by the automated evaluation system because these students are not given nearly as much attention to the understanding of dominant cultural constructs (Holstein & Doroudi, 2022). This matter is particularly problematic in intercultural or multilingual learning situations, when AI tools cannot identify the variety of student learning patterns and thus provide unfair final grades and waste the chance to tailor academic guidance (Pedro et al., 2019). Unless accompanied by purposeful adjustments like fairness audits, bias reduction methods, and participatory AI design that integrates multiple educational stakeholders, these AI systems are likely to accentuate educational disparities and negate the ethos of fairness and inclusion as the backbone of ethical educational practices with teaching and learning (Eynon, 2023).

Algorithmic Bias and Its Sources in Educational AI

Algorithms with pervasive biases emerged as one of the highly disputable ethical issues in the domain of educational systems where one can apply the capabilities of AI. The frequently recurring training data sets can influence the judgment of AI systems because they encourage bias and social discrimination at school, as indicated by various scholars (Muoneke, 2025). The algorithms that build AI learn the patterns on the basis of the previous academic results, demographics, and behavioral patterns/pointers—data sets that are at best biased and often not representative of diverse student bodies. Both articles by (Doroudi, 2024), highlighted that the algorithmic decision based on this kind of biased data is unfair and often has negative effects on groups that are marginalized the most, which would fall under the traditionally underserved socio-economic, racial, or linguistic background. More recent research has proven that the AI technologies that are used to conduct admissions, predict performance, and identify students reproduce the inequalities built into these systems by giving an advantage to students with the same kinds of data about them that historically favored groups already had (Hintz, 2024). These findings show that it is essential to challenge the origins of bias in algorithm AI systems in education and create models of data equity.

The absence of contextually and culturally relevant data in AI training models is a significant contributor to bias. The social and cultural complexity of varied student groups may not be reflected by AI systems primarily based on data integrated into uniform, affluent, or metropolitan educational contexts, particularly in developing countries (Rasool et al., 2025). As one of the

examples, predictive analytics systems aimed at determining at-risk students sort out learners belonging to marginalized or rural groups as being academically at risk due to biased covariates but not because of the actual academic proficiency (Whittaker et al., 2019). This problem is complicated by the fact that culturally sensitive variables are not included and the educational standards of the West are overrepresented in the development of AI models (Holstein & Doroudi, 2021). A recent study in South Asian and African settings established that AI-based decisions in education are disproportionately excluding students of underrepresented populations because they did not appear in their training data sets (Rahmanipur & Mohammadi, 2025). These differences raise the necessity to apply context-specific models of AI that would correspond to the challenges of different educational contexts.

Moreover, algorithmic bias in education decisions is inherently enabled by the design (or structure) of AI algorithms. A significant number of AI systems applied in education use predictive models based on the quantitative values (grades, attendance rates, standardized test scores, etc.) and drop or disregard the qualitative indicators (creativity, critical thinking, emotional resilience, etc.) in students ((Oyetade & Zuva, 2025). This rank uses numerical data, which replicates current inequalities since students with marginalized backgrounds will not necessarily play equal roles on traditional academic measures because of the socio-economic forces facing them or their cultural differences in being able to learn (Khosravi et al., 2022). Another study by (Eynon, 2023) further claimed that the potential outcome of these algorithmic practices is an extension of existing educational disparities, as it adheres to skewed assumptions constituted in data and decision rules. Also (Williamson & Eynon, 2020), noted that unless the system of education utilizes AI has very clear fairness audits and bias mitigation procedures, then it would exacerbate education inequalities instead of addressing them. These lessons demonstrate that it is worth structuring AI models introducing a wide variety of performance metrics and fair decision-making rules.

Research Hypotheses

1. There is a significant presence of algorithmic bias in AI-driven educational tools that affects classroom decision-making processes.
2. There is a significant negative impact of AI-driven biased decision-making on students' academic opportunities, recognition, and educational equity.
3. There is a significant association between implementing fairness audits, inclusive design strategies, and reduced bias in AI-based educational decision-making.

Methodology

Research Design

The current study was based on the use of a cross-sectional survey research procedure and a semi-quantitative research design in determining the extent to which the problem of algorithmic bias in the use of AI education technology influences the way teachers view the equality in the field of education in the classroom. It was determined that the design was appropriate at a specific period of time in order to gather the views, experiences, and fears amongst tutors concerning the use of AI in the classroom. The questionnaire structure gave the researchers an opportunity to evaluate the responses of the participants with the assistance of a semi-quantitative approach, and in addition, it provided the researchers with the opportunity to describe it. The choice of this form could be explained by the fact that the same methodology was used in the studies of similar educational environments, and the similar concepts, in general, such as technology coverage, social justice of

algorithms, and educational inequality in general, were also addressed. Other factors that influenced the decision include the fact that the study was cross-sectional, thus allowing them to collect information about a large scope of the population in a small amount of time.

Population

The participants of this study were the school and university teachers in the Punjab province of Pakistan. The purpose of selecting the teachers was that they were selected in both the public and private schools because it can cover the different institutions of organizations and grade levels of teachers. The study population consisted of teachers whose level of AI use and awareness differed, which allowed us to get a whole picture of their attitude toward the question of algorithmic bias and how it affects educational fairness. The preference of teachers as the key group was also vital because they have direct contacts with the AI meanings employed in educational administration, student evaluation, and educational aid, and, therefore, their observations are somewhat involved in the explicitly held research questions. Moreover, educators are in a central position to not only deal with ethical issues associated with the adoption of AI but also support and promote fair approaches to that in their schools or universities. The study focused on this group because it sought to extract as much information as possible on the people most exposed to the effects of AI decision making in the learning institution.

Sample Size

The sample of the study was 270 teachers. Research conducted in the educational technology research industry had taken a sample size; in a similar manner, it was researched in this study, and it was assumed that it was adequate and that realistic statistical analyses like Pearson correlation, independent sample t-tests, and regression analysis could be carried out through it. The researchers have taken care to ensure sample-wide coverage of academic background, years of experience teaching, and years of knowledge on AI. There were both teachers in school and university, as they were the target of the study in ensuring the study of perception of the AI fairness to the various educational setups. It was a good sample that helped the study achieve statistical validity, although it was not impractical to the research resources and time available to implement the study. What is more, the extensive and diverse sample enhanced the capacity of the outcomes of the research to be generalized to a broader educational context in the province of Punjab. It also increased the odds of acquiring subtle opinions concerning the positivity or negativity of algorithms as the relationship of AI applications, bias, and fairness appeared.

Sampling Method

The snowball sampling technique was applied in the study to select the participants. The first thing to do is to select a group of teachers that the researchers are familiar with and invite them to take the survey. These educators were then requested to share the survey link with other educators and colleagues in the profession, making it easier to include more individuals in the survey. The reason why this sampling technique was selected is that it is effective and convenient to use, as it will be possible to contact educators who work in public and private institutions in varied cities of Punjab.

Data Analysis

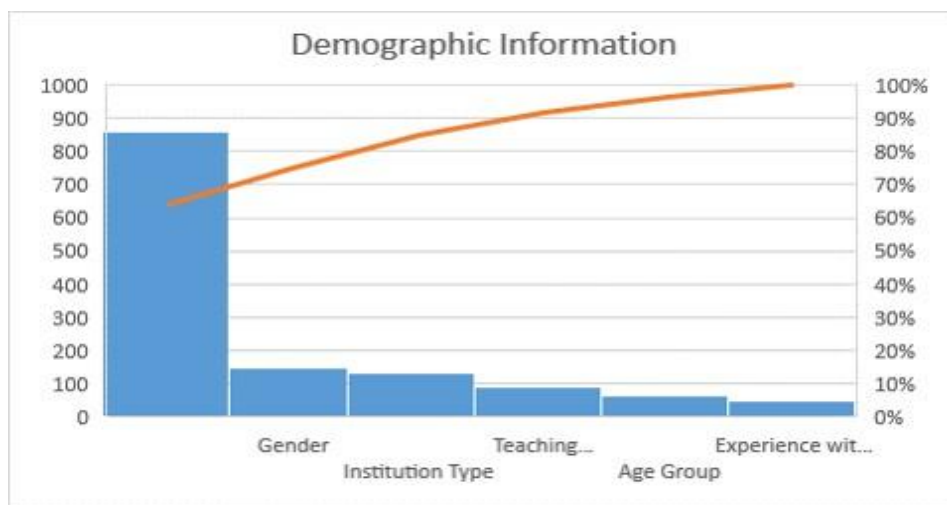
Data entry and subsequent analysis were run on the Statistical Package for Social Sciences (SPSS) version 26 that was introduced using answers provided by 270 respondents. As a way of summarizing the demographic aspect of the respondents and other important study variables, descriptive statistics were used, including frequencies, percentages, means, and standard

deviations. Inferential statistical tools were used to examine the hypotheses of the study. The Pearson correlation was performed to understand the relationship between the algorithm bias in educational AI tools and the perceptions of the teachers with aspects of educational equity. The independent samples t-test was utilized to establish if there were any significant differences between the school and university teacher perceptions of algorithmic bias. In addition, a basic linear regression operation was carried out to verify the accuracy of the predictive effect of the AI experience of teachers on their perceptions of algorithmic bias. Any statistical test will also be done on a $p < 0.05$ level of significance to either accept or reject a hypothesis. This well-developed plan of analysis guaranteed very rigorous, trustworthy, and statistically significant analysis of the research questions.

Table 1: Demographic Analysis for Respondents (N = 270)

Variable	Category	Frequency (n)	Percentage (%)
Gender	Male	150	55.6
	Female	120	44.4
Age Group	21–30 years	65	24.1
	31–40 years	110	40.7
	41–50 years	70	25.9
	51 years and above	25	9.3
Institution Type	School	135	50.0
	University	135	50.0
Teaching Experience	Less than 5 years	90	33.3
	5–10 years	100	37.0
	More than 10 years	80	29.6
Experience with AI Tools	No Experience	50	18.5
	1–2 years	100	37.0
	3–5 years	80	29.6
	More than 5 years	40	14.8

Figure 1:



Based on the demographic data of the 270 people who participated in the research, it was noted that there was a gender-balanced population with 55.6 percent males and 44.4 percent women. Most of the respondents (40.7%) ranged between 31 and 40 years, followed by 25.9 percent of 41 to 50 years, 24.1 percent of 21 to 30 years, and 9.3 percent of 51 years and above. There was a balanced ratio of school teachers (50.0%) and university teachers (50.0%) to guarantee representation of all institutional views. As to the experience in terms of teaching, 37.0% got between 5 and 10 years of service, 33.3% less than 5 years, and 29.6% more than 10 years of experience in teaching. As far as the use of the AI tools is concerned, 37.0% stated having 1-2 years of experience, 29.6% stated having 3-5 years of experience with AI tools, and 18.5% and 14.8% stated having no experience and more than 5 years' experience, respectively. This distribution shows a rich demographic sample, which is quite representative and comprehensive to the analysis of the perceptions and biases about AI in the educational field.

Table 2: Correlation Analysis for Hypothesis H1

Variables	Mean	SD	1	2
1. Algorithmic Bias Perception	3.45	0.68	1	
2. Educational Equity Perception	3.71	0.72	.512**	1

The correlation done during the analysis of Hypothesis H1 showed that the strong perception of algorithmic bias revealed that there existed a strong linkage between the perception of educational equity and the perception of equality in the classroom between teachers. The Pearson product-moment correlation coefficient was calculated to estimate an R value of 0.512 with $p < 0.01$, which indicated the relationship between the two variables to be moderate and positive. This observation suggests that the higher the percentages of the increase of the issue of algorithmic bias in AI education tools, the more the teachers are concerned with the absence of equity in education. In other words, the sense of bias behind the teachers in an AI-driven decision-making apparatus would bias the teachers to question the fairness and inclusiveness of an AI.

Technology in the classroom. The positive correlation highlights the necessity to make an investment in the idea of the algorithmic fairness where the principles of educational justice and equality have to be maintained in the artificial intelligence-aided educational environment. The given findings are justified by the earlier research concerning the potential of AI bias to influence the dependence and the feeling of justice of teachers and an educational system.

Figure 2:

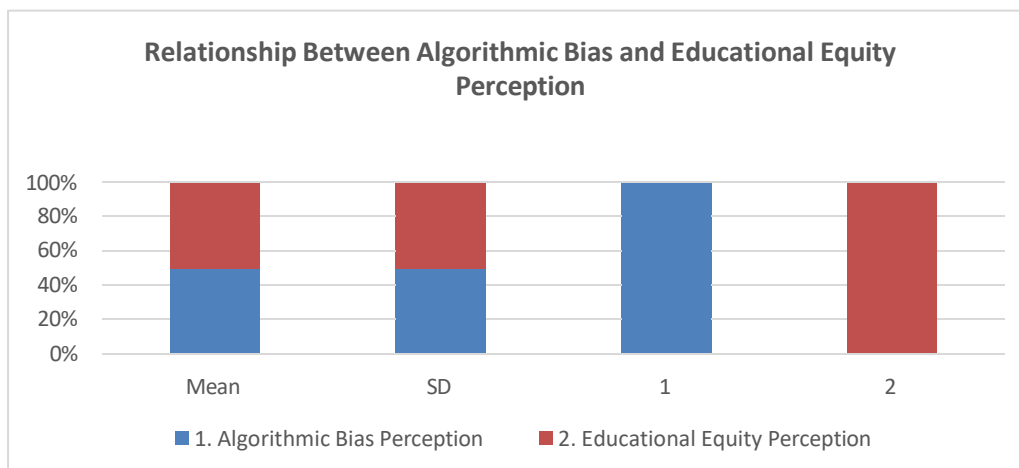
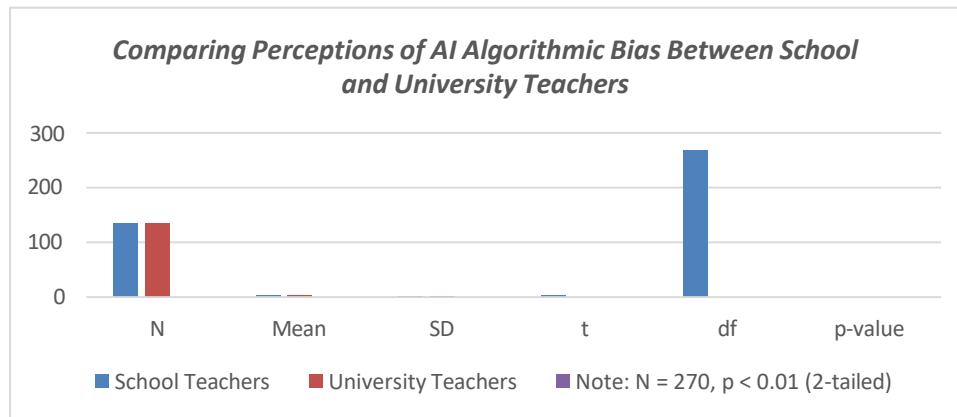


Table 3: t-Test Analysis for Hypothesis H2

Group	N	Mean	SD	t	df	p-value
School Teachers	135	3.52	0.66	2.87	268	0.004**
University Teachers	135	3.33	0.70			

An independent samples t-test was used to determine the difference between the perception of AI algorithmic bias of the school teachers and the university teachers. The statistical test demonstrated the statistically significant difference, $t(268) = 2.87$, $p = 0.004$, where the perceptions of algorithmic bias were higher in the respective domain among school teachers ($M = 3.52$, $SD = 0.66$) than university teachers ($M = 3.33$, $SD = 0.70$). The given finding helps to state that school teachers tend to think that bias will manifest itself in AI educational tools more readily than higher education teachers do.

Figure 3:



The finding indicates that the teaching environment and the roles could affect the experience and assessment of AI-based systems by educators in learning places. Such contrasts point to the necessity of approaching the issues of AI fairness in academic contexts with some focus on the educational context.

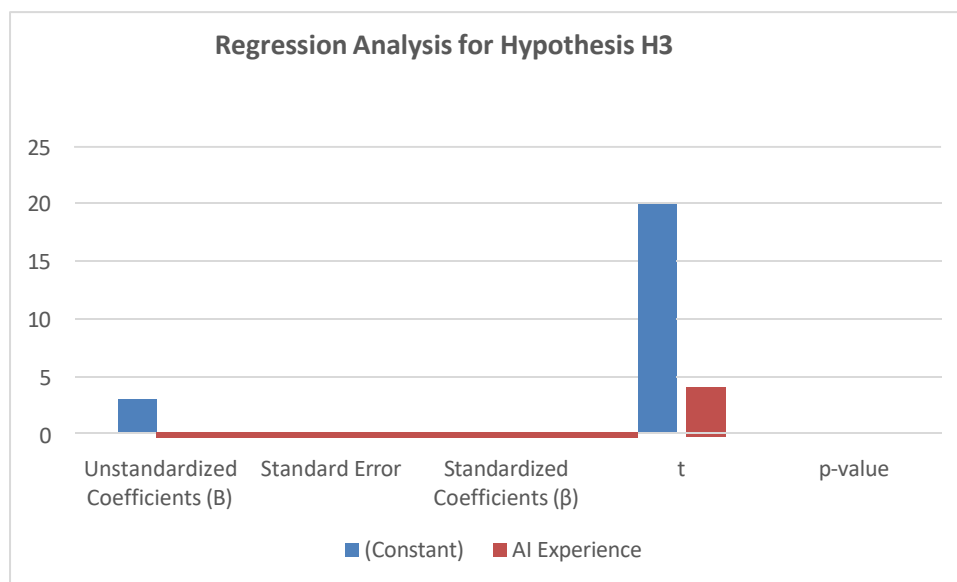
Table 4: Regression Analysis for Hypothesis H3

Model	Unstandardized Coefficients (B)	Standard Error	Standardized Coefficients (β)	t	p-value
(Constant)	2.98	0.15	—	19.87	0.000**
AI Experience	0.21	0.05	0.322	4.24	0.000**

A linear regression was used to perform a simple regression analysis because the first goal was to determine whether the experience with AI tools significantly predetermined an attitude towards the presence of algorithmic bias in the educational state as a predictor and perceptions of the presence and sources of algorithmic bias in the educational state as a dependent variable. These results revealed that AI experience served as a high positive predictor of bias perception, $B = 0.21$, $t = 4.24$, $p = 0.000$. It implies that the perception of algorithm bias was raised by 0.21 units for each increment of 1 unit of AI experience. The constant 2.98 denotes the perceived score that would be predicted in the case that AI experience is at zero. The model was statistically significant, and it

accounted for a meaningful percentage of variance in bias perception scores. The two variables do have a moderate positive relationship with the standardized coefficient ($= 0.322$). These results indicate that the higher the experience of the teachers in working with AI educational tools was, the more likely they were to find the existence of algorithmic bias in classroom experiences, which aligns with the hypothesis.

Figure 4:



Discussion

The results of the present study put a dangerous emphasis on the increasing concern of teachers about the responsibility of the algorithmic bias within AI-based educational tools and its consequences on education equity. The correlational analysis proved the existence of a moderate positive association between algorithmic bias perception and worry about fairness in AI-mediated decision-making. This implies that the more teachers became aware of the existence of bias in AI tools, the more they felt concerned about equal access to academic opportunities. The relationship supports the discussion found in the previous literature that when unregulated, AI systems can become one more factor perpetrating inequalities and dichotomies and making the practices of pedagogy less inclusive (Binns, 2022). It becomes evident that the propensity of teachers to such biases affects their trust in the ethical use of AI in classrooms, which must be taken critically in the AI policy frameworks in education. Moreover, an independent samples t-test found a significant difference in perceptions of school and university teachers regarding algorithms and their possible biases. The level of concern about bias in education AI tools was higher among school teachers than it was among their university counterparts.

Such a difference can be explained by the fact that the roles, responsibilities, and classroom management experiences of the teachers working in the school setting are significantly different, as AI-based systems are frequently applied to such sensitive practices as monitoring student behavior and predicting their performance. This finding is consistent with the earlier studies showing that perceptions of AI technologies depend on the environment educators operate in, and front-line users tend to be less concerned about the ethical issues than administrators and supervisors (Holstein & Doroudi, 2021). The findings warrant the fact that AI governance policies should be adapted to the unique challenges and expectations of teachers at different levels of education. The regression analysis also confirmed the hypothesis stating that the experience of

using AI tools is a strong predictor of algorithmic bias perceptions in the teachers. The educators who had more familiarity with the educational systems grounded on AI were found to describe most instances when bias was built into these technologies. The observation can echo other literature pieces that found an improved level of familiarity with AI systems tended toward identifying both technical and ethical weaknesses in algorithmic decisions in educators (Eynon, 2022). The predictive power of the AI experience in constructing bias perception implies that professional development plans that emphasize the literacy of AI have the potential to be a significant determinant in ensuring responsible employment of AI in the educational field. Empowering teachers with the skills necessary to critically assess the use of AI tools in education will not only enhance the education attainment of students but will also guarantee that the ethical rule of thumb like fairness and inclusivity are observed.

Finally, demographic analysis revealed a high and well-balanced representativeness of the respondents, including the survey of different ages, gender categories, teaching experience, and the degree of AI awareness. Such distribution of the populations helped form a complete picture of the varying levels of attitudes towards AI bias among the various learning environments. Both the balanced representation of the school and university teacher also served to enhance validity of the comparative analysis between types of institutions. As it is AI perception studies that present similar demographic inclusivity, previous researchers suggest that the use of such demography will result in more accurate localization of contextual issues that, in turn, will lead to developing AI implementation strategies that will be not only fair but also responsive to the individual needs of every stakeholder in the educational environment (Mohammadi & Maghsoudi, 2025). The sample wide range thereby made it possible to carry out a balanced examination of the issues that bother the teacher consequently underlining the importance of having broad based AI policy frameworks in the education sector of Pakistan.

Conclusion

To sum up, the analysis conducted in this paper has shown a critical correlation between the perceptions of algorithmic bias in AI educational tools and the teachers concern regarding educational equity, which have shown remarkable correlations that are of critical concern as the study progresses regarding the AI-driven classrooms. The results showed that the teachers who appeared to be under higher threats of algorithmic bias were more likely to report worries regarding fairness and inclusivity, and the bias was observed more in the school teachers than in the university teachers. Also, experience using AI instruments revealed itself as a strong forecaster of educator's awareness about bias, showing that the more educators are exposed to those systems, the sharper the discussion of ethical implications develops. These findings confirm the findings of current international studies concerning the dangers of the uncontrolled use of AI in the field of education and the lack of proper training of the educational workers of all educational establishments, transparency and openness of AI governance frameworks, and justice in AI design to protect against the threat of discrimination in the educational process. By tackling these issues, educational centers in Pakistan and other countries can make sure that AI technologies will become the tool of betterment and not the method of exclusion and inequality.

Recommendations

1. Implement AI Literacy Programs: Schools and universities need to conduct periodic workshops to improve the knowledge of educators about the possibility of AI tools, their parameters, and how they can negatively affect the learners.

2. Conduct Fairness Audits: AI-used educational tools are to perform regular fairness and bias audits to discover and eliminate discrimination in the algorithms prior to its use.
3. Establish Transparent AI Policies: Institutional policies on how AI is ethically used in the classroom should be made easily understandable and accountable by putting in place transparent measures.
4. Through Pedagogy and Equity by involving teachers at all levels of teaching: Involve the educators in the development and testing of AI tools, making sure they are standardised in accordance with pedagogical values and equality principles.
5. Adjust AI Practices to Various Contexts: AI practices must be tailored to the various contexts: AI practices must be tailored to various contexts, such as school and university settings individually, given that these settings have different challenges and teaching dynamics.
6. Place Value on Inclusive Data Practices: AI developers need to employ various, proximate, and relevant data sets to reduce the cultural and demographic biases that may be inherent in them.
7. Enforce Ethical Oversight Committees: Form AI ethics review board in schools to control the choice, application and assessment of AI systems.
8. Advocate Ongoing Investigations: Advocate continued empirical research in the field of AI bias and educational equity to track changing issues and knowledge in terms of beneficial integration of AI.

Future Implications

The results of the research also have a significant implication on the development of future education practices, the making of policies, and the development of the AI systems within the academic settings. With the further AI-driven transformation of the teaching and learning process, educational institutions, technology developers, and policymakers must focus on fairness, inclusivity, and transparency in the AI-based decision-making tools. Future attempts have to underline the development of AI in a context-sensitive and culturally encompassing way, particularly within the educational environment in which diversity is the norm, as is the case in Pakistan. Moreover, the paper suggests that it would be helpful to establish long-term teacher education to foster the knowledge of AI and ethical preference in teachers so that they would be in a position to assess AI technologies in their classrooms. This study also promotes future empirical studies on long-term social, psychological, and educational effects of AI-mediated learning, which should result in making future developments lead to the improvement but not the memorization of equity and justice in educational settings.

References

1. Barnes, E., & Hutson, J. (2024). Navigating the complexities of ai: The critical role of interpretability and explainability in ensuring transparency and trust. *International Journal of Multidisciplinary and Current Educational Research*, 6(3).
2. Binns, R. (2018). Fairness in machine learning: Lessons from political philosophy. Conference on fairness, accountability and transparency,
3. Binns, R. (2022). Human Judgment in algorithmic loops: Individual justice and automated decision-making. *Regulation & governance*, 16(1), 197-211.
4. Binns, R. (2024). If the Difference Principle Won't Make a Real Difference in Algorithmic Fairness, What Will? Response to 'Rawlsian Algorithmic Fairness and a Missing Aggregation Property of the Difference Principle'. *Philosophy & Technology*, 37(4), 119.

5. Binns, R., & Kirkham, R. (2021). How could equality and data protection law shape AI fairness for people with disabilities? *ACM Transactions on Accessible Computing (TACCESS)*, 14(3), 1-32.
6. Bulathwela, S., Pérez-Ortiz, M., Holloway, C., Cukurova, M., & Shawe-Taylor, J. (2024). Artificial intelligence alone will not democratise education: On educational inequality, techno-solutionism and inclusive tools. *Sustainability*, 16(2), 781.
7. Chinta, S. V., Wang, Z., Yin, Z., Hoang, N., Gonzalez, M., Quy, T. L., & Zhang, W. (2024). FairAIED: Navigating fairness, bias, and ethics in educational AI applications. *arXiv preprint arXiv:2407.18745*.
8. Doroudi, S. (2024). On the paradigms of learning analytics: Machine learning meets epistemology. *Computers and education: artificial intelligence*, 6, 100192.
9. Dwivedi, Y. K., Kshetri, N., Hughes, L., Slade, E. L., Jeyaraj, A., Kar, A. K., Baabdullah, A. M., Koohang, A., Raghavan, V., & Ahuja, M. (2023). Opinion Paper:—So what if ChatGPT wrote it? Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy. *International journal of information management*, 71, 102642.
10. Eynon, R. (2022). Datafication and the role of schooling: Challenging the status quo. In *Learning to live with datafication* (pp. 17-34). Routledge.
11. Eynon, R. (2023). Algorithmic bias and discrimination through digitalisation in education: A socio-technical view. In *World Yearbook of Education 2024* (pp. 245-260). Routledge.
12. Hintz, A. (2024). AI, big data and bias: governing datafication through a data justice lens. In *Handbook of Media and Communication Governance* (pp. 526-537). Edward Elgar Publishing.
13. Holstein, K., & Doroudi, S. (2019). Fairness and equity in learning analytics systems (FairLAK). Companion proceedings of the ninth international learning analytics & knowledge conference (LAK 2019),
14. Holstein, K., & Doroudi, S. (2021). Equity and artificial intelligence in education: will "aied" amplify or alleviate inequities in education? *arXiv preprint arXiv:2104.12920*.
15. Holstein, K., & Doroudi, S. (2022). Equity and artificial intelligence in education. In *The ethics of artificial intelligence in education* (pp. 151-173). Routledge.
16. Holstein, K., Wortman Vaughan, J., Daumé III, H., Dudik, M., & Wallach, H. (2019). Improving fairness in machine learning systems: What do industry practitioners need? Proceedings of the 2019 CHI conference on human factors in computing systems,
17. Khosravi, H., Shum, S. B., Chen, G., Conati, C., Tsai, Y.-S., Kay, J., Knight, S., Martinez-Maldonado, R., Sadiq, S., & Gašević, D. (2022). Explainable artificial intelligence in education. *Computers and education: artificial intelligence*, 3, 100074.
18. Khosravi, I. (2025). Towards sustainable agriculture in Iran using a machine learning-driven crop mapping framework. *European Journal of Remote Sensing*, 58(1), 2490787.
19. Mohammadi, A., & Maghsoudi, M. (2025). Bridging perspectives on artificial intelligence: a comparative analysis of hopes and concerns in developed and developing countries. *AI & SOCIETY*, 1-22.
20. Muoneke, M. (2025). From Crisis to Insights: Leveraging Data Analytics to Shape Pandemic-Era Educational Technology Innovations. Available at SSRN 5134393.
21. Ntoutsis, E., Fafalios, P., Gadiraju, U., Iosifidis, V., Nejd, W., Vidal, M. E., Ruggieri, S., Turini, F., Papadopoulos, S., & Krasanakis, E. (2020). Bias in data-driven artificial intelligence systems—An introductory survey. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 10(3), e1356.
22. Oyetade, K., & Zuva, T. (2025). Advancing Equitable Education with Inclusive AI to Mitigate Bias and Enhance Teacher Literacy. *Educational Process: International Journal*, 14, e2025087.

23. Pedro, F., Subosa, M., Rivas, A., & Valverde, P. (2019). Artificial intelligence in education: Challenges and opportunities for sustainable development.
24. Rahmanipur, A., & Mohammadi, S. (2025). Examining the Role of Artificial Intelligence in Enhancing Educational Equity: A Systematic Review. The First International Conference on Artificial Intelligence, Tehran, Iran,
25. Rasool, S., Lodhi, H. A., & Hussain, I. (2025). AI and the Future of Learning: Personalization, Equity, and Ethical Challenges. *Journal of Social Signs Review*, 3(4), 247-259.
26. Shen, H., Deng, W. H., Chattopadhyay, A., Wu, Z. S., Wang, X., & Zhu, H. (2021). Value cards: An educational toolkit for teaching social impacts of machine learning through deliberation. Proceedings of the 2021 ACM conference on fairness, accountability, and transparency,
27. Stephenson, B., & Harvey, A. (2022). Student equity in the age of AI-enabled assessment: Towards a politics of inclusion. In *Assessment for Inclusion in Higher Education* (pp. 120-130). Routledge.
28. Swist, T., & Gulson, K. N. (2023). School choice algorithms: data infrastructures, automation, and inequality. *Postdigital Science and Education*, 5(1), 152-170.
29. Veale, M., Van Kleek, M., & Binns, R. (2018). Fairness and accountability design needs for algorithmic support in high-stakes public sector decision-making. Proceedings of the 2018 chi conference on human factors in computing systems,
30. Whittaker, M., Alper, M., Bennett, C. L., Hendren, S., Kaziunas, L., Mills, M., Morris, M. R., Rankin, J., Rogers, E., & Salas, M. (2019). Disability, bias, and AI. *AI Now Institute*, 8, 11.
31. Williamson, B., & Eynon, R. (2020). Historical threads, missing links, and future directions in AI in education. In (Vol. 45, pp. 223-235): Taylor & Francis.
32. Williamson, B., Komljenovic, J., & Gulson, K. N. (2023). Introduction: Digitalisation of education in the era of algorithms, automation and artificial intelligence. In *World yearbook of education 2024* (pp. 1-19). Routledge.
33. Yeung, K., & Harkens, A. (2023). How do "technical" design-choices made when building algorithmic decision-making tools for criminal justice authorities create constitutional dangers? Part II. *arXiv preprint arXiv:2301.04715*.