INTERNATIONAL JOURNAL OF ADVANCE RESEARCH, IDEAS AND INNOVATIONS IN TECHNOLOGY

ISSN: 2454-132X

Impact Factor: 6.078

(Volume 10, Issue 4 - V10I4-1204)

Available online at: https://www.ijariit.com

Biases in Ai-Driven Educational Tools

Aanya Yaduvanshi
aanya2106007@gmail.com
Delhi Public School R.K Puram, New Delhi

ABSTRACT

This paper explores the prevalence of biases in AI-driven educational technologies, examining their sources, manifestations, and impacts on diverse student populations. It proposes a framework for identifying and mitigating biases, emphasizing the need for transparent and inclusive educational technologies. The global market value of AI in education is projected to increase by over 45% CAGR(Compound Annual Growth Rate) between 2022 and 2030- highlighting the vast potential of this growing industry. As a responsible society, we must ensure the ethical reproduction of Artificial Intelligence in a sensitive domain like education, where the potential is boundless however the risks are equally potent. This paper aims to set the foundation of AI as an assistive technology and not as a replacement for traditional teaching. The research highlights the use of Deep Learning and Natural Language Processing in Curriculum Design or Pedagogy planning, focusing on critical thinking and not only static learning. While other recent works have merely reviewed one use case, we have categorized EdTech tools into three learning pathways: Student-Supporting, Teacher-Supporting, and System-Supporting. By theoretical and empirical analysis, we systematically break down two technologies Assistive Grading Technology and Chatbot Personalized learning, and consequently explore every impact on the stakeholders: students, teachers, and the educational community. Furthermore, we emphasize the need for a safe regulatory framework, much like the one currently existing, to ensure there is no misuse of data collected. The laws surrounding the fragility of the information sourced and the usage of data mining for further replication should be strict, with appropriate legal consequences for breaching the same.

Keywords: *Artificial Intelligence, Bias, Education, Assistive Grading, Data Bias, Algorithmic Bias, Bias Mitigation, Machine Learning*

1. INTRODUCTION

With the rampant increase in AI-driven technology in education tools, through assisted teaching, grading, and adaptive learning systems, there comes a need to review current existing biases and subsequently ethical implications for the future. From integrating Generative AI into building classroom curriculums to using NLM for automated grading essays, AI must be approached not only in terms of expansion but also from fairness and bias perspectives. This paper emphasizes the diverse implications of these biases affecting the education community and calls for transparent and inclusive educational technologies. A concerted effort by a collaboration of interdisciplinary stakeholders, including policymakers, technologists, educationists, and students will be needed for the efficient renewal of policies around transparency, accountability, and equitability. Using a mixed-methods approach, combining model analysis and qualitative insights from interviews, the paper presents a recommended framework for identifying and mitigating these biases, specifically three: data, algorithmic, and interaction bias. Valuable insights from the literature review lead us to provide tangible bias mitigation techniques. This paper aims to re-evaluate the potential biases in AI-driven educational tools and demand an ethical system for consideration before the introduction of them into educational communities. For teachers and administrators, the challenge put forth would be to ensure the safe integration of this technology into teaching practices and to analyze its impact on student engagement and neural learning pathways. They would be required to weigh all aspects of this new change and then suggest tangible improvements for the same. The effects of generative AI must be carefully examined and discussed, considering learners' over-reliance on the same. Since AI systems are increasingly being used in diverse sectors, even other than education, healthcare, law, and science, they are particularly prone to amplifying existing societal biases and inequities. Thus, the paper provides effective solutions to ethical problems created due to the introduction of any AI-powered tool into the field of education. Through the paper, we analyze the threat of AI-enabled influence operations and outline steps to be taken before language models are used globally. We hope that the paper spurs in-depth research on challenges stemming from bias like disinformation and equity breakdown, which are currently beyond the scope of this research.

2. Literature Review

Previous works have recorded implications on different target populations and categorized them on the basis of potential negative impact and practicality. Implicit bias in grading platforms, specifically due to linguistic variations and irregular perspectives, has been noted to reflect inherent bias towards the stereotypical idea of what constitutes as 'normal' [1]. Other frameworks which have recorded bias include Intelligent Tutoring systems [2]. In accordance with the concept of sustainable education, ITS proves to be a highly efficient model of providing personalized learning tools to students, catered to individuals' unique needs and preferences.
With respect to the 'pedagogical troika' of education- teaching, learning, and assessment- language technologies have a massive potential to enable equity [3]. However, with prevalence of biases, they may derail the actual purpose of education, and thus must be combatted at both a grassroot and global level for leveraging the advantages, and not the risks. A stepwise theoretical framework, like that proposed by the European Commission [4] incorporates collaborative internal review, pilot testing and opportunities to back out in case of ethical concerns. Such systems holistically account for all pillars of privacy, data governance, fairness, trustworthiness, accountability and inclusivity. Subsequently, there will be a need to analyze other relevant models from fields combining educational psychology, computer science, and ethics.

Specifically for curriculum designing, certain principles have already been successfully integrated in technology-driven learning tools [5].This particular goal has been at the forefront of all AI-based learning tools subsequently, where outsourcing of automation is worked upon, and not of critical thinking [6].

This is supported by the theory of Universal Design for Learning, which is an educational framework that promotes equitable and flexible methods to accommodate the needs of diverse learners, making it more personalized [7]. Flexibility in curriculum design allows for adaptation to different educational contexts and student needs [8]. However, while that is the ideal goal, inherent sources of bias like gender, race, language all have the capacity to replicate themselves in AI, systematically and persistently. Thus, manifesting biases in diverse ways in the curriculum. Examples include but are not limited to biased recommendations for course materials, unequal representation of diverse perspectives, and biased assessments of student performance.

These biases can disadvantage underrepresented groups and reinforce existing inequalities. Since most algorithms are trained predominantly on historic Western-centric datasets, they frequently present a skewed representation in their algorithms and interaction, thus showcasing a lack of cultural sensitivity which can alienate marginalized communities[9]. In fact, even without usage of EdTech solutions, there exists problems in areas like Standardized reading achievement, where curriculum bias leads to discrepancies in grade equivalent scores between different reading curricula and tests for a single curriculum [10].

Owing to previous studies and frameworks, we already have pedagogies which incorporate equity learning, and subsequently aim to reduce the unintended harmful consequences on learners [11]. Historically too, educators have advocated for fairness in education, but now this too must be implemented in the new-age of innovation with education technology. Important to note is the difference in approaches of scientific learning as compared to theoretical and humanistic focuses. For both, there are scalable methods, which are culturally relevant, feasible and holistically modern. Finally, Sustainable AI curriculum planning should be informed by self-determination theory and include content, product, process, and praxis approaches to ensure continuous refinement and relevance [12].

Before, during, and after integration of technology into Pedagogy Planning and testing for K-12 curriculums especially, there needs to be consensus between parents and teachers about the potential harmful consequences. Thus, it is imperative for all stakeholders to be aware of the varying degrees of implications in a learner's future.

With the wide prevalence of edtech tools, there is a large amount of data being collected for analysis, expanding potential in data mining and potential misuse. Specifically in the case of surveillance capitalism, students report feeling anxious when being made aware of how their privacy is breached when their digital activities are being tracked [13]. Thus, there arises a need for interdisciplinary collaboration between Algorithm Development and Information Security, by strict legislations being enforced through laws and legislations.

To ensure that there is no risk of noncompliance with these privacy regulations, there must be strict establishment of clear standards and guidelines for AI development, thus ensuring adherence to best practices for fairness and inclusivity. These standards address data diversity, algorithmic transparency, and accountability[14]. Additionally, legislation can enforce anti-discrimination laws, ensuring AI systems do not disproportionately impact marginalized groups. Along with that, independent audits can provide objective assessments, ensuring AI systems operate fairly and equitably.

3. Research methodology and Data Collection

We conducted semi-structured interviews with experts from different fields, classified into two: subject-knowledge experts and industry experts. We use the term industry experts to refer to someone who is currently working in the development of AI-driven educational tools, and is actively involved in the product's growth. Subject-knowledge experts refer to educationists, research scholars who are professionals in fields pertaining to Edtech, AI, DeepLearning networks, and NLP.

We interviewed 3 experts from all different organizations, which raged from startups with 25 employees to corporations with more than 5000 employees. The experts ranged from Post Doctoral research associates to the Co-Founder of a DeepTech company. We recruited interviewees by emailing potential candidates at organizations and reaching out via LinkedIn. In some cases, we contacted individuals who forwarded us to experts within their organization. The recruitment strategy does not introduce any inherent bias, since the participants range from all across India. Our research goal with these interviews was to help clarify specific technologies like Assisted Grading Platforms, outline challenges, and characterize pathways for the efficient roll-out of AI-driven educational tools.

3.1. Interview questions and finding insights

We began each interview by introducing the purpose of the interview and subsequent research: to experts' analysis and views on certain pertinent questions regarding AI and its influence in Edtech, specifically certain tools like AI Assisted Grading and Chatbot Learning, and any challenges they face. Each interview lasted from 30 minutes to 1 hour. As most interviewees were outside of New Delhi, we conducted interviews over Google Meet.
We asked several open-ended questions and encouraged interviewees to describe their particular field of interest and specialization, and further asked follow-up questions on domain-specific expertise. In each interview, we aimed to learn the following:

- What is their role in the company's workflow of AI-Edtech solutions?
- What technology does their product use?
- How is their company structured to work towards Bias Mitigation in their specific product?
- Is there a dedicated team or role responsible for bias detection in their organization?
- How much emphasis is placed on fairness, transparency and equity in development of the model?
- Do they assess and address potential biases in the datasets used for training AI models?
- Is there a feedback mechanism for users to report perceived biases in AI outputs?
- Are there measures in place to address AI biases that might impact students' learning experiences?
- Is there a concern that AI biases could affect the credibility and adoption of EdTech solutions into society?
- Do teachers have resources to understand and mitigate AI biases in educational tools they use?
- How do they see the industry progressing in the upcoming years, especially with the prevalence of Generative AI in learning?
- Are stakeholders (e.g., students, teachers) involved in the data collection process to ensure it is representative?
- What are the recurring challenges they face in their work?

Other than these questions, there were also technology-specific questions. For example, during the interview on Assisted Grading Platforms, questions like ' How does your technology ensure that it does not replicate the same biases enforced by teachers during grading? '. Similarly, for the Semantic Search in the Chatbot personalized learning tool, there were questions like "How does their organization ensure that the vast amount of personalized data collected is not misused in cases of data harvesting?".
During the interviews we took extensive notes, and then transcribed them further after in-depth analysis and comparison.

4. Discovery/Results

4.1. Teacher-supporting: Assistive Grading Platforms by Smartail.AI

Smartail.AI, a prominent deeptech company specializes in AI and its subsets, such as computer vision, natural language processing (NLP), and deep learning algorithms. Their product DeepGrade focuses on solving the use case of assessment, since it is an important tool in measuring students' understanding level. Through their product, they try to help teachers by empowering them in the teaching process and subsequently the remedial learning for the students.

The interviewee, Mr. Aslam Sherieff is the co-founder of Smartail.AI. He aided us in answering specific questions about the technology used in DeepGrade, as well as general questions about the structure of bias mitigation in their organization. We have obtained his consent for naming him and the company.

4.1.1. Algorithm, Labeling Method and Dataset Curation

Previously, the technology was based on a model prediction system; this magnified the bias since there is already wrong data being fed into the bias prediction model. Instead, they use a unique method which incorporates tools like patent-filed Knowledge Graph. They incorporate a method called Labeling, wherein different datasets are curated into multiple levels of data.This process ensures that each label is correct so that no erroneous results are produced . Since even one wrong data might give a biased result, it is very important to rigorously test the model against each data in every step, by evaluating it through multiple iterations. Separate production data is obtained from high-end GPUs from companies like Amazon and Nvidia. By regular training and measuring bias outcomes on a daily basis, they reach a certain desired accuracy level. Although it is highly impractical to reach 100% accuracy level, they try to maximize accuracy and reliability by minimizing bias, as per international standards.

'All grading technologies have slight variations, however we are improving with time. We ensure that the algorithm is tuned efficiently to successfully compare machine and human graded responses simultaneously.'

The logic behind assessment is that each student's answers are not compared and graded, but instead assessed separately, making their technology responsible and fair. Without any sentiment and emotion in play, the answers are purely looked at from technical perspectives as prescribed in the answer key attached. Thus, the source of each answer is the same as per the rubric input. When the handwritten answer sheets are scanned and uploaded as PDF documents, it is processed in the backend: extraction of handwritten text, grading using Natural Language Processing, matching with answer key or value points. The model also takes into account the uniqueness of each student, and is thus able to apply and infer the meaning of children's answers.

4.1.2. AI in Education

Traditional assessment methods like Periodic/Formative/Summative Assessments are all being slowly augmented with technology-based solutions: replacement of pen-and-paper tests to digital models and multiple-choice objective questionnaires. DeepGrade aims to enhance educational outcomes by automating grading of handwritten answer papers. Their entire application has been collaboratively developed by teachers, educational institutes, and students. They have worked with schools for two years at no cost in order to cocreate the model, by taking feedback on training datasets, explaining variations and measuring appropriate level of accuracy.

'Our technology maps student responses, differentiating between memory-based questions and those requiring higher-order thinking skills (HOTS) and critical thinking. In fact, this approach aligns with the CBSE board's emphasis on practical application, complimenting a significant transformation in the educational sector in India.'

4.1.3. Addressing Learning Gaps

After the assessment is done, the AI analyzes the missing points of the answer and identifies it as a particular learning gap, and a potential pattern for the rest of the class. After several assessments, proper understanding flaws are generated, and the teacher is tasked with creating targeted remedial plans to establish expected learning outcomes. For instance, if only 50% of expectations are met then the teaching would incorporate learning with experimental approach, AR-VR, Exploratory work, autonomous learning etc.

4.1.4. Enhancing Teacher Efficiency

By automating the grading process, the technology significantly reduces the time teachers spend on assessments, allowing them to focus more on analytics and strategy development for reteaching and reassessment.

'This shift enables teachers to implement autonomous learning strategies, addressing knowledge gaps early and improving overall educational outcomes. The system has been piloted in over 100 schools and colleges, demonstrating objective and reliable assessment capabilities.'

4.2. Student- supporting: Chatbot personalized learning

4.2.1 Advanced Video Generation in Enhancing Personalized Learning

The company employs advanced AI technologies to create educational videos tailored to specific student needs. This approach leverages Large Language Models (LLMs) and various AI tools to generate content dynamically, ensuring a personalized and efficient learning experience.

4.2.2 Personalized Video Generation

When a student requests to see a specific concept, such as a projectile motion in the z-axis, the model can streamline the learning process. Instead of requiring the student to sift through an entire lecture, LLMs process the student's prompt, extracting relevant information from internet sources and conducting backend analysis. The prompt is then sent to models like Stable Diffusion or

DALL-E to generate visual content, which is combined into a 15-20 second video clip using Whisper library in Python for transcription and text-to-speech APIs for audio narration.
'The generated video is saved in a library, tagged and titled with the user's prompt. This enables quick retrieval for any subsequent similar queries from other users. If a new query deviates significantly from the saved prompt (by more than 70-80%), the system generates a new unique video to ensure accuracy and relevance.'

4.2.3 Integrated Learning Analytics

In order to help carry out appropriate diagnostic and remedial measures, similar to Deepgrade, this technology integrates learning pace analysis on the basis of tests. It generates holistic reports, suggesting focus areas and providing recommendations. Additionally, it carries out doubt-solving but only up to a certain intermediate level, since one cannot fully rely on LLMs to solve complex doubts. Moreover, if there are only ready-to-consume answers, the whole ideology of learning and improving gets nullified, in accordance with the theory of practical learning.

4.2.4 Integration of Rule-Based and Intelligent Deep Learning Models in Chatbots

The company developed and applied intelligent models like bird transformer models, LSTM(Long Short-term memory) networks, which emphasize the major role of technology in the edtech industry.
The company works towards the 3 biases in all areas of work.
Data Bias: While building any AI model application, the engineers firstly describe the use cases and explicitly only work on solving those. Whatever data they utilize, has been nurtured and feature engineering by formation and shaping to get a specific desired output.
'If there are N-number of datapoints, and you only require 10, then you will specifically choose those 10. Otherwise you make the model unnecessarily heavy and overburdened. There is a variance-bias tradeoff, you cannot complicate or oversimplify the model since it will either not learn anything or will overfit respectively.

Algorithmic Bias: During development of the algorithm, sometimes due to the human-business side biases may occur. For instance, in classification algorithms, if it detects humans in the frame, the model learns only detection and in steps after in application, the logic is built up on top of the bias during development. However, there are techniques involving intelligent systems which are able to avoid that.
Interaction Bias: Similar to the Turing test, after development of a model, it competes to give an outcome which is near-human. The entire challenge is to provide results which are human-based but not human-driven. However, in recommendation engines like music applications, even during the initial onboarding, the app takes your preferences and an inherent bias is introduced. Similarly, in educational platforms, the user chooses a particular recommendation without knowing the backend generation. It may choose a particular algorithm from the frontend and unknowingly for future references, the model recommends outcomes from the user's preferences, on top of the business logic.

4.2.5 Semantic Search

The semantic search tool refers to an umbrella for the whole application, which enables the user to navigate to each part of the app just by interaction with the bot. This is particularly useful for users who are not familiar with the app, and unaware of most features, the automatic navigation aiding the user experience.
'Educational applications are often vast, with numerous features that users may be either aware of or unaware of. The integration of LLMs and semantic search helps users navigate these extensive applications more efficiently. This technology not only improves user experience but also ensures that all features of the application are accessible through simple interactions.'

4.3. Bias Mitigation strategies

Here, we present a step-wide checklist for product developers to follow to ensure both Data Bias, and Algorithmic Bias do not heavily influence the results.

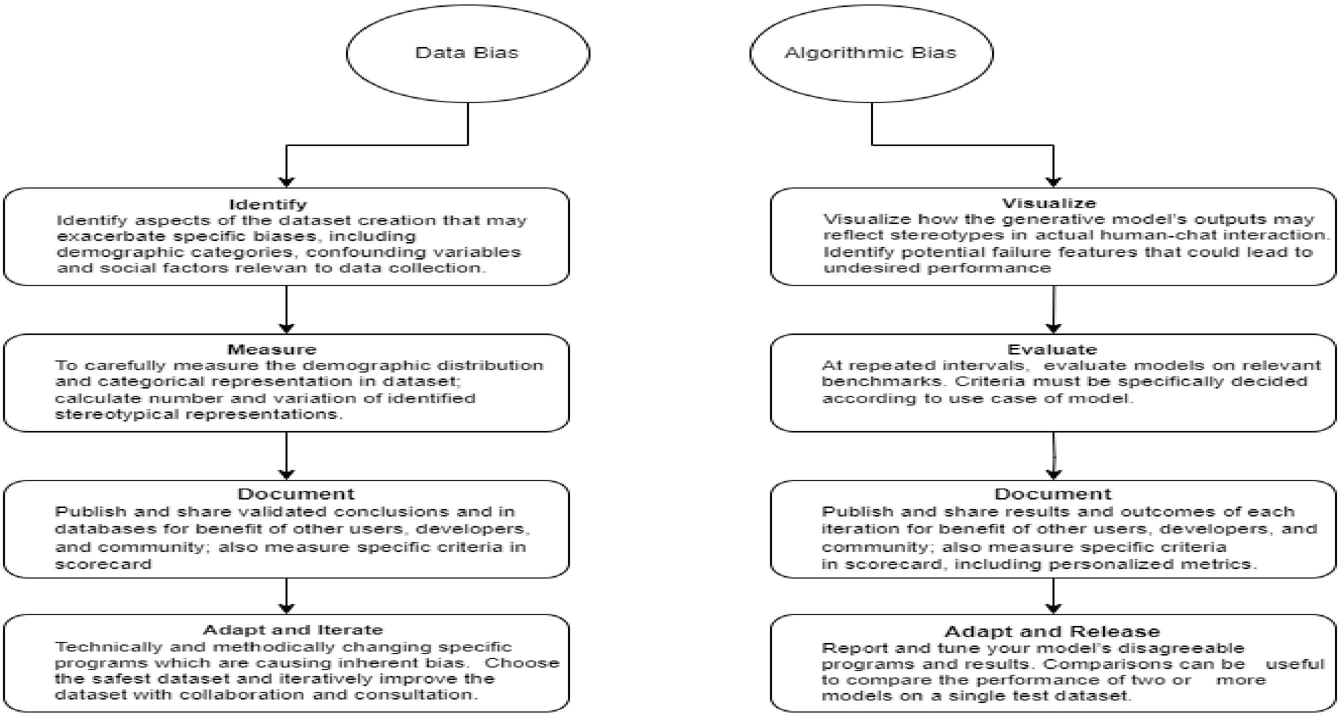

Figure [1] Bias Mitigation Framework

A concise and complete strategy for detecting, documenting and removing Data and Algorithmic Bias from Artificial Intelligence models

Bias is complex, both in detection and mitigation. To ensure that each step is accounted for, there are several standard metrics used for determining features including: Disparate Impact and Equalized odds test. Apart from those, there are certain non-parametric cohort analyses which essentially refers to employing statistical tests that do not assume the data fits into one fixed probability distribution, such as the normal bell curve. There are many different nonparametric statistical tests, but for scope of this research, we recommend using the McNemar test, a paired nominal test.

4.4 Stakeholder analysis

After deep conversation and quantitative analysis, the research states that AI's potential to enhance teaching and learning is tempered by concerns regarding biases and their implications for the credibility and equitable adoption of these technologies. This research focuses on the perspectives of key stakeholders—teachers, students, and the broader educational community—to elucidate the complexities surrounding AI integration in education.

4.4.1 Teachers: Skepticism and Adaptation

Educators are pivotal in the adoption and effective utilization of EdTech solutions. According to industry experts, many teachers perceive AI as an assistive technology rather than a replacement for traditional pedagogical methods. This perception is crucial for maintaining AI's credibility in educational contexts. However, by surveying, teachers express widespread skepticism regarding the reliability and transparency of AI models used for grading and assessment. The apprehension among educators stems from the potential biases inherent in AI systems, which could unfairly impact student evaluations.

'The integration of AI necessitates substantial adaptation, especially among senior teachers who may resist technologies perceived as encroaching on their professional roles. To address these concerns, it is essential to position AI as a tool for enhancing educational efficiency and engagement, rather than supplanting the educator's role.' Comprehensive professional development and training programs are imperative to enable teachers to understand and effectively utilize AI technologies.

4.4.2 Students: Dependency and Learning Efficacy

The introduction of AI in K-12 education significantly affects students, particularly concerning the potential for increased dependency on AI-driven solutions like ChatGPT. Although AI offers personalized learning for students, there is a risk of hampering their independent learning capability. For example, tools which generate customized video content based on specific queries, exemplify how AI can tailor educational material to individual needs.

However, these solutions must complement rather than replace traditional learning methods to preserve the integrity of the educational process, ensuring holistic development instead of static learning, as mentioned in the Introduction.

'ChatGPT has made , especially students, handicapped; having replaced interaction via search engines, there is evidence that Google Search significantly went down with introduction of Generative AI. Unless it is very specifically utilized as assistive technology, there may be negative outcomes for students in K12. But you can't exactly stop it. The only solution ahead is building awareness.'
The use of AI for personalized feedback and doubt resolution can significantly enhance learning efficiency. Nonetheless, it is crucial to maintain a balance to prevent over-reliance on AI, and ensure balanced learning. Educational strategies should include educating students on the limitations and responsible use of AI technologies.

4.4.3 Educational Community: Awareness and Equity

The broader educational community, encompassing parents, policymakers, and researchers, must address the ethical implications of AI in education. The digital transformation spurred by the COVID-19 pandemic has increased awareness of AI, yet this awareness is uneven across different regions and socio-economic groups, and can only be bridged by conscious efforts.
'After extensive surveys and pilot testing for their services, we concluded that in tier 3 and 4 cities, there is barely resistance to new technologies in EdTech. However, in tier 1 cities, they are supposedly busy with administration so they aren't very enthusiastic about integration of AI into assessment grading.
AI biases pose a significant challenge to the equitable adoption of EdTech solutions. Unaddressed biases in AI models can perpetuate existing inequalities and introduce new ones. Ensuring that AI systems are transparent, accountable, and rigorously tested is crucial for minimizing biases. In order to foster complete and unbiased understanding of AI''s capabilities, there must be active creation of AI literacy programs for all stakeholders involved.
Workshops and tutorials at the school level, involving both students and parents, can demystify AI and highlight its potential benefits and risks. More active teacher training is needed, to showcase how their role in remedial classes is more needed, but reduced in automated tasks like grading.

4.5 Safety Measures in Data Collection and Utilization

The Information Technology (IT) laws enacted in 2019 in India provide stringent guidelines regarding data collection and user privacy. According to the Personal Data Protection Bill(2019), sensitive personal data includes information related to financial data, biometric data, caste, religious or political beliefs, and other categories specified by the government. It also clearly mentions, 'no person shall collect this sensitive personal data without explicit effective consent from the data subject'. Thus, bringing the concept of consent strengthens the rules surrounding data collection.
As per the interviews, the concept of data protection extends to both explicit and implicit data collection practices. For instance, implicit user profiling, which can infer sensitive information from non-direct sources such as the type of mobile phone used, is equally restricted.
For educational purposes, it is permissible to collect less personalized data, such as school age or grade, since it directly helps in providing personalized and appropriate recommendations, provided that it does not lead to sensitive inferences. To safeguard data integrity and privacy, it is crucial to employ robust security provisions for company data use . Applicable to any company size and any industry, ISO 20000 gives the company the methodology and the framework to prove that their company follows best practices, thus ensuring that data is securely transmitted and stored. This certification ensures that any data entering or leaving the organization is encrypted, preventing unauthorized access or breaches. They ensure user trust, privacy regulations and ethical code of conduct set by international standards.
'Maintaining data sanctity is a core feature of our data management strategy. Encrypted databases are employed to store personal information, ensuring that even internal developers or unauthorized personnel cannot easily access or decipher this data.'

5. Discussion and implications

This paper has highlighted technologies including Automated Essay Grading, which in conformation with data from interviews, is mostly recommended for subjects like English, thus ruling out linguistic and dialectical differences. On the contrary, the Automated Grading and Feedback system of Smartail's DeepGrade assessment is specifically focused on subjects with clearly specified marking schemes like in Biology, ruling out areas which have scope of perspective. Both have their own algorithms, which either predict or are based on DeepLearning and NLP. Due to their dependence on these machine learning models, concerns of bias in learning technologies are on the rise.
We identify three specific stages of bias detection and subsequent mitigation - task definition, data curation, and model training.
Bias in models utilizing Artificial Intelligence (AI), Deep Learning, and Natural Language Processing (NLP) is a multifaceted issue. Bias, in this context, refers to systematic favoritism or discrimination that AI systems may exhibit towards certain groups or individuals based on characteristics such as race, gender, or socioeconomic status.
Among the various types of biases, three main categories significantly impact these processes: Data Bias, Algorithmic Bias, and Interaction Bias. Data bias occurs when the training and testing datasets are unrepresentative or incomplete, leading to inherent biases in the model's outputs. Algorithmic Bias refers to the inaccurate and biased assumptions and metrics in the algorithms' design and functioning that are reflected in the machine's outputs. Finally, Interaction Bias, or User bias occurs when during interaction

with the AI model, people introduce their own conscious or unconscious biases and prejudices into the system, specifically in the case of interactive chatbots.

Educational tools leveraging AI have been classified into three categories on the basis of their use cases.

Student supporting tools are those which directly aid the student's learning process by resource providing, intelligent tutoring systems(ITS), direct interaction via bots etc.

On the other hand, teacher supporting, or facing, tools are those which are responding to teacher's requirements such as automated grading platforms, curriculum designing, remedial concept teaching etc.

System supporting tools are wide-ranging, including diagnostic capabilities, educational data mining for resource allocation, administrative tasks etc.

Majority of these tools provide personalized, one-on-one, 24x7 assistance, and thus have massive potential to revolutionize traditional education. However, the implications need to be considered during introduction and implementation in real-world learning pathways, as outlined in Table 1.

	STUDENT SUPPORTING	TEACHER SUPPORTING	SYSTEM SUPPORTING
DATA BIAS	• Skewed and mismatched pedagogical recommendations • Inaccurate and outdated resource allocation	• Preliminary student analysis flawed and non-holistic • Fragmented and primitive curriculum design	• Failure in diagnosing modern learning difficulties (underrepresentation in dataset) • Wrong categorization of learning materials and methods
ALGORITHMIC BIAS	• Hallucinations leading to distorted tutoring assistance • Amplify stereotypes in classrooms due to logical error	• Unfair evaluation in automated grading systems • Unable to personalise to each student's requirements	• Potential security risk in educational data mining • Unfair highlighting of students needing additional learning support
INTERACTION BIAS	• Replication of students' inherent biases into algorithm • Compromise student's ability to think critically and creatively by overtaking independent learning	• Chatbot teaching assistors marginalize already discriminated students • Unable to adapt to student's requests and needs	• Globalise existing inequalities by repeated iteration of errors • Obstacle in productivity due to constant sources of contradiction

Table[1]

Classification of Bias and EdTech tools with cross-referencing implications

6. Challenges

AI and Machine Learning have made promises that are proving to be difficult to keep. With great advancement comes feasibility in implementation and most importantly, acceptance in the 21st-century society. As discussed in the qualitative interviews, many K12 schools are reluctant to shift to more technology for learning, teaching, and assessment. Questions arise on practical and feasible integration into institutions and simultaneously ROI(return-on-investment). All stakeholders must derive a consensus on which risks to prioritize; and what problems we are willing to solve with technology, with their probability of success. There is no reason to hand over every single global challenge to technology simply for the sake of novice advancement. The need of the hour is to assess every aspect of Responsibility, Transparency, and Accountability. All sensitive tasks about the personal data of users must be entrusted to professionally trusted and reliable platforms. Otherwise, concerns regarding the privacy of data collection, data mining, and surveillance capitalism may abound. Culturally relevant pedagogies must be curated, keeping in mind historical systematic discrimination against marginalized communities. Training datasets and testing models must be optimized to result in minimum Algorithmic Bias so that there are no unforeseen effects during interaction with the model. Most datasets are skewed toward Western ideologies, and model pilot testing has only been done in English. This can unconsciously create culturally biased outcomes, affecting the adoption of AI in certain sensitive regions having linguistic and ethnic differences. To ensure that there are no such undesired consequences, the model must undergo several iterations and improvements to remove user biases and replications in the model during the interaction. As highlighted in Table 1, the implications of inaccurate and biased results are wide-ranging. Thus, appropriate solutions must be found and steps taken to ensure that these challenges do not outweigh the massive benefits of AI in education.

7. Conclusions

Through interviews and research, it is clear that AI should only be used as a potential assistant to teaching, and not as a replacement. Nonetheless, it is quite evident that AI holds great promise in revolutionizing fields pertaining to Intelligent Tutoring Systems, Curriculum Designing, Assistive Grading Technology, including combinations of the aforementioned. Successful integration of these technologies requires careful analysis and discourse among all stakeholders on matters of ethics, fairness, trust and responsibility. Specifically, it must be ensured that the three biases involved in creation of an AI-based technology(Data Bias, Algorithmic Bias, and Interaction Bias), is taken care of by a unique team focused on mitigating the same. Concerns surrounding widespread adoption in Tier 1, 2 and 3 cities must be met with awareness programs, literacy initiatives and training bootcamps, both for teachers and students. The only way forward to maximizing the potential of Machine Learning tools in K12 schools is by increasing teacher productivity and enhancing student learning outcomes. The integration of rule-based intelligent models and advanced deep learning techniques,such as that mentioned by the interviewees, is reshaping the functionality of educational applications. Leveraging open-source models like LLAMA 2 for constructing these systems enhances features like semantic search capabilities, resulting in more efficient user interaction and ease. Provisions of detailed performance analysis, alongwith effective remedial strategies, fosters a conducive environment to independent and practical learning, in line with Sustainable Development Goal 4(Quality Education) as outlined by the United Nations[16].

REFERENCES

[1] Khan, S. The Ethical Imperative: Addressing Bias and Discrimination in AI-Driven Education. Social Sciences Spectrum 2023, 2 (1), 89–96.

[2] Lin, C.-C.; Huang, A. Y. Q.; Lu, O. H. T. Artificial Intelligence in Intelligent Tutoring Systems toward Sustainable Education: A Systematic Review. Smart Learning Environments 2023, 10 (1). https://doi.org/10.1186/s40561-023-00260-y.

[3] Gordon, E. W.; Rajagopalan, K. Assessment for Teaching and Learning, Not Just Accountability. Palgrave Macmillan US eBooks 2016, 9–34. https://doi.org/10.1057/9781137519962_2.

[4] Directorate-General for Education, Y. Ethical Guidelines on the Use of Artificial Intelligence (AI) and Data in Teaching and Learning for Educators; Publications Office of the European Union: LU, 2022

[5] Matiisen, T.; Oliver, A.; Cohen, T.; Schulman, J. Teacher-Student Curriculum Learning. arXiv.org. https://doi.org/10.48550/arXiv.1707.00183.

[6] Mayfield, E.; Madaio, M.; Prabhumoye, S.; Gerritsen, D.; McLaughlin, B.; Dixon-Román, E.; Black, A. W. Equity beyond Bias in Language Technologies for Education. Proceedings of the Fourteenth Workshop on Innovative Use of NLP for Building Educational Applications 2019. https://doi.org/10.18653/v1/w19-4446.

[7] Wu, Y. Integrating Generative AI in Education: How ChatGPT Brings Challenges for Future Learning and Teaching. Journal of Advanced Research in Education 2023, 2 (4), 6–10.

[8] Chiu, T. K. F. A Holistic Approach to the Design of Artificial Intelligence (AI) Education for K-12 Schools. TechTrends 2021, 65 (5), 796–807. https://doi.org/10.1007/s11528-021-00637-1.

[9] Khan, S. The Ethical Imperative: Addressing Bias and Discrimination in AI-Driven Education. Social Sciences Spectrum 2023, 2 (1), 89–96.

[10] Jenkins, J. R.; Pany, D. Curriculum Biases in Reading Achievement Tests. Journal of Reading Behavior 1978, 10 (4), 345–357. https://doi.org/10.1080/10862967809547288.

[11] Mayfield, E.; Madaio, M.; Prabhumoye, S.; Gerritsen, D.; McLaughlin, B.; Dixon-Román, E.; Black, A. W. Equity beyond Bias in Language Technologies for Education. Proceedings of the Fourteenth Workshop on Innovative Use of NLP for Building Educational Applications 2019. https://doi.org/10.18653/v1/w19-4446.

[12] Chiu, T. K. F.; Chai, C. Sustainable Curriculum Planning for Artificial Intelligence Education: A Self-Determination Theory Perspective. Sustainability 2020, 12 (14), 5568. https://doi.org/10.3390/su12145568.

[13] Kumi-Yeboah, A.; Kim, Y.; Yankson, B.; Aikins, S. K.; Yvonne Appiah Dadson. Diverse Students' Perspectives on Privacy and Technology Integration in Higher Education. British Journal of Educational Technology 2023, 54 (6), 1671–1692. https://doi.org/10.1111/bjet.13386.

[14]Smith, G.; Rustagi, I. Mitigating Bias in Artificial Intelligence an Equity Fluent Leadership Playbook; 2020. https://haas.berkeley.edu/wp-content/uploads/UCB_Playbook_R10_V2_spreads2.pdf.

[15] United Nations. Goal 4: Quality Education. The Global Goals. https://www.globalgoals.org/goals/4-quality-education/.

.