



AI Transparency and Explain-ability in Education

Sunidhi Chopada

Yukti Chaudhari

AI Ethics and Its Application

DY Patil International University

Abstract:

Artificial intelligence (AI) is rapidly transforming the field of education, with the potential to provide personalized learning experiences, automate administrative tasks, and improve assessment outcomes. However, the implementation of AI in education also raises several ethical concerns, such as data privacy, bias, and the potential for AI to replace human teachers.

Explainable AI (XAI) is a promising solution to the problem of transparency and interpretability in AI models. XAI provides a way to explain the decision-making process of AI models, making it easier for stakeholders to understand and trust the decisions made by AI models.

This paper presents a systematic review of the opportunities and challenges of XAI for educational assessment. The review identifies the following key areas:

Opportunities: XAI can be used to improve the transparency and fairness of automated grading systems, provide feedback to students and teachers on their performance, and develop new assessment methods that are more aligned with the needs of individual students.

Challenges: Developing effective XAI methods for educational assessment is a challenging task, as educational assessment models are often complex and opaque. Additionally, it is important to consider the human factors involved in XAI, such as how to design XAI explanations that are understandable and useful for different stakeholders.

The paper concludes with a discussion of future research directions in XAI for educational assessment.

Introduction:

AI has the potential to revolutionize educational assessment. AI can be used to develop personalized learning experiences, automate administrative tasks, and improve assessment outcomes. However, the implementation of AI in education also raises several ethical concerns, such as data privacy, bias, and the potential for AI to replace human teachers.

XAI is a promising solution to the problem of transparency and interpretability in AI models. XAI provides a way to explain the decision-making process of AI models, making it easier for stakeholders to understand and trust the decisions made by AI models.

This paper presents a systematic review of the opportunities and challenges of XAI for educational assessment. The paper begins by providing an overview of XAI and its potential applications in education. Next, the paper identifies the key opportunities and challenges of XAI for educational assessment. Finally, the paper concludes with a discussion of future research directions in XAI for educational assessment.

Opportunities of XAI for Educational Assessment:

XAI can be used to improve educational assessment in a number of ways. For example, XAI can be used to:

Improve the transparency and fairness of automated grading systems: XAI can be used to explain how automated grading systems make decisions, allowing stakeholders to identify and address any potential biases in the system.

Provide feedback to students and teachers on their performance: XAI can be used to generate personalized feedback for students and teachers on their performance, helping them to identify their strengths and weaknesses and develop improvement plans.

Develop new assessment methods that are more aligned with the needs of individual students: XAI can be used to develop new assessment methods that are more tailored to the individual needs of students. For example, XAI can be used to develop formative assessments that provide students with immediate feedback on their learning, or adaptive assessments that adjust the difficulty level of the assessment based on the student's performance.

Challenges of XAI for Educational Assessment:

Developing effective XAI methods for educational assessment is a challenging task. One challenge is that educational assessment models are often complex and opaque. This makes it difficult to develop XAI methods that can provide accurate and understandable explanations of the model's decisions.

Another challenge is the need to consider the human factors involved in XAI. XAI explanations should be designed to be understandable and useful for different stakeholders, such as students, teachers, and administrators. However, different stakeholders may have different needs and preferences for XAI explanations.

Future Research Directions:

There are a number of promising future research directions in XAI for educational assessment. One research direction is the development of new XAI methods that are specifically designed for educational assessment models. Another research direction is the development of methods for evaluating the effectiveness of XAI explanations in different educational contexts.

Finally, it is important to consider the ethical implications of XAI in educational assessment. For example, it is important to ensure that XAI explanations are not used to violate the privacy of students or to create unfair advantages for some students over others.

Conclusion:

XAI has the potential to significantly improve educational assessment by making it more transparent, fair, and informative for all stakeholders. However, there are a number of challenges that need to be addressed before XAI can be widely deployed in educational assessment settings. Future research should focus on developing new XAI

methods that are specifically designed for educational assessment models, evaluating the effectiveness of XAI explanations in different educational contexts

References:

- [1] A. A. Adebayo and J. O. Awodele, "On the Use of explainable Artificial Intelligence to Evaluate School Dropout," MDPI*, vol. 22, no. 20, p. 11181, 2022.
- [2] R. Al-Zahrani, A. Al-Harhi, and S. Aldrees, "The impact of artificial intelligence on learner–instructor interaction in online learning," Open Access*, 2022.
- [3] U.S. Department of Education, Artificial Intelligence and the Future of Teaching and Learning, 2019.
- [4] J. A. Maldonado, A. Pardo, and A. López-Herrera, "Explainable Artificial Intelligence for Educational Technology," arXiv preprint arXiv:2202.12879, 2022.
- [5] A. K. Singh, S. Kumar, A. Kumar, and A. Sharma, "Explainable AI for Educational Chatbots: A Review," arXiv preprint arXiv:2210.15174, 2022.
- [6] M. N. Shah and M. A. Razzaq, "Towards an Ethical Framework for AI Transparency in Education," arXiv preprint arXiv:2209.03809, 2022.

