An Analysis on the application of Explainable AI (XAI) in Course Assessment for Effective Teaching and Learning Process

Dr.THOMAS FELDMAN

Vellore Institute of Technology,

Chennai

Abstract

The Challenges in Teaching and Learning process generally includes curriculum design, content delivery, Evaluation / assessment, students' behavior analysis and improvement in terms of learnability and mental ability of the students. The goals of TLP are: (i) improve the learning curve of the students and (ii) providing the students the opportunities to meet out the real world requirements. However, there are challenges in improvement of learning ability of students due to the inability to take the feedbacks given after the evaluation in a positive manner; unable to find improvement tactics from the regular feedbacks. Thus the need of the hour is adopting novel methodologies to meet the personalization based improvement requirements as each student is unique.

Keywords : Teaching and Learning Process (TLP), Explainable AI (XAI), Personalized Feedback, Assessment, Feedback

1. Introduction

The new paradigm in AI is Explainable Artificial Intelligence (XAI) which integrates a set of methods and processes that assures in delivering trustable results generated by the Machine Learning algorithms. The XAI provides higher level of accuracy and precision in the generated outcomes to get a highly empowered decision making process.

The tools which are applied in teaching and learning process are nowadays equipped with AI in order to meet out the dynamic needs and trustworthy solutions in a timely manner. The previous AI based teaching learning tools are like black box in which the internal processing involved to arrive at a result are completely unknown and cannot be concluded how a particular result has been arrived. However, with the invention of XAI, this issue has been resolved.

However, there are some research challenges and opportunities involved on the application of XAI in the evaluation process. (i) to improve the learning ability and performance improvement of the students, it is essential to have a learning analytics platform that embeds XAI with the more effective learning method (ii) to deliver an intelligent and explainable feedback to the students and the corresponding actions to be taken in a data driven manner (iii) The teachers or management who is the data owner of the feedbacks of evaluation should be able to append recommendations or actions by means of active leaning method to reassure the stakeholders such as students and parents that there are no biases which may affect the actions to be taken.

As XAI has transparency and trustworthiness in its outcome generation, it will definitely help in predictive analytics of performance level of the students by means of self regulated data driven analytics with personalized recommendations based on the analytics to improve the learning curve of the students.

1.1 Research Objectives

This research work has the following objectives:

- To improve the learning curve of the students.

- To provide a trust worthy and an un-bias feedback to the students for improvement.

- To generate meaningful feedback reports based on ML based analytics in XAI to the stakeholders in the TLP.

1.2 XAI – Introduction

Explainable Artificial Intelligence (XAI) integrates a set of methods and processes that assures in delivering trustable results generated by the Machine Learning algorithms. It provides higher level of accuracy and precision in the generated outcomes to get a highly empowered decision making process. **XAI** helps in justifying why a model has made a particular choice. It makes the decision transparent by providing explanations that can be understood by non-experts.

The **XAI** is focused on the development of multiple systems by addressing challenge problems in two areas:

- machine learning problems to classify events of interest on heterogeneous, data; and

- machine learning problems to construct decision policies for an autonomous system to perform a variety of simulated missions.

- To handle the above challenging areas, **classification and reinforcement learning** are applied to perform **intelligence analysis on autonomous systems**.
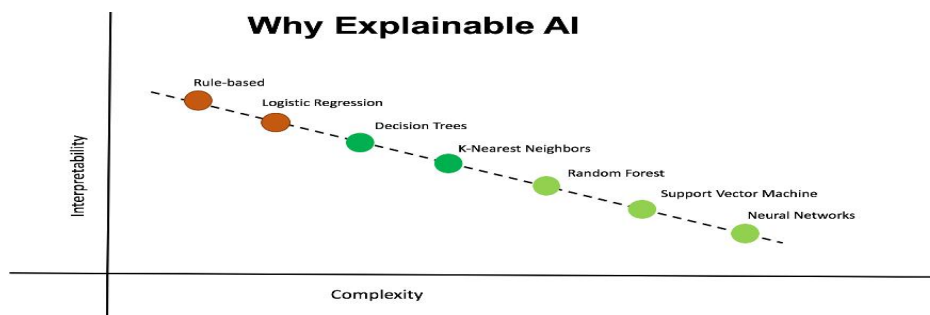
Figure 1 – Need for XAI in real time systems

(Ref.: Client Solutions - https://analyttica.com/understanding-shap-xai-through-leaps/)

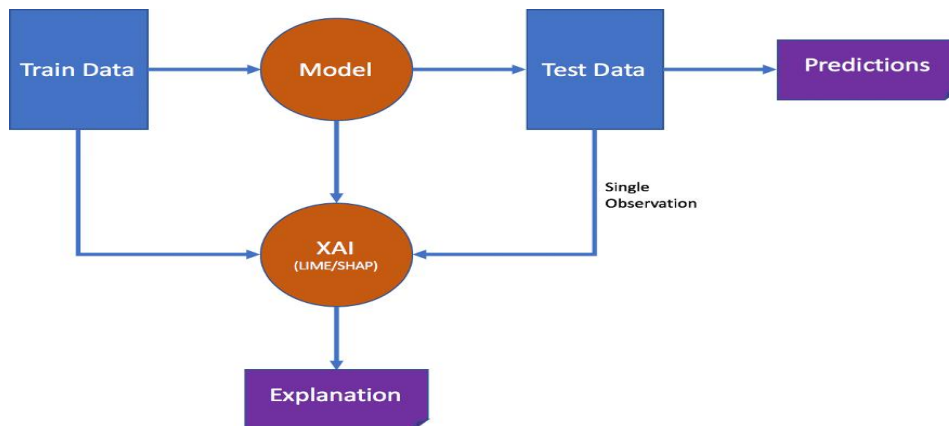**1.3 Generic workflow of XAI**

Figure 2 – Generic Workflow of XAI

The Explainability can be of the following categories: Model-Specific , Model-Agnostic and Model-Centric Application of XAI in this research work has utilized the **transparency** and **trustworthiness** in its outcome generation. it helps in the predictive analytics of performance level of the students by means of **self regulated data driven analytics** with recommendations based on the analytics.

Here, the Data-Centric Explainability is applied. It helps to understand the nature of the data and provides consistent results. It is well appropriate for solving the TLP problems as data plays a crucial role in building prediction and classical modeling, it is necessary to understand the algorithm's behavior concerning the given dataset. \

The Approaches applied in this work are:

➢ Data Profiling - the process of examining, analyzing, and creating useful summaries of data.

➢ Monitoring Data-Drifts – monitoring statistical properties of the target variable, which the model is trying to predict, change over time

➢ Data-Adversarial – Identification of adversarial attacks in machine learning as they attempt to disrupt ML models with deceptive input data and prediction requests.

2. **Literature Review**

Saeed and Omlin(2023) have performed an extensive meta survey on the application of XAI and its challenges and opportunities. Their review indicated the research directions in this new AI

paradigm. Their study also examined the difference between explainability and interpretability. They have provided a development life cycle of XAI based system such as Design, Development and Deployment. They have also examined the trustworthiness of XAI in real time domains implementation. Their study concluded on two major areas: distinction between interpretability and explainability and challenges and research directions on the application of XAI.

Haqueet.al(2023) have provided a synthesis on XAI literature and the future research directions on using XAI. Their analysis indicated four dimensions of XAI: completeness, accuracy, format and current aspects. Also, they have identified the effects of XAI in applying it for real time domains in terms of usability, fairness, transparency, trust and understandability.

Clement et.al(2023)have introduced a novel framework namely XAIR and provided the methods and tools used in implementing XAI. They have also proposed an SDLC model with life cycle phases Requirement analysis, Design, Implementation, Testing and Deployment.

 Nwakanma et.al (2023) have provided a review on the application of XAI in intrusion detection and mitigation techniques using XAI in intelligent connected vehicles. They have identified the characteristics of XAI such as transparency, repeatability and trustworthiness can be applied to provide a safe interconnected intelligent vehicles. They have provided a review on XAI models for intrusion detection and other open research issues found in XAI.

 Khosravi et.al (2022) have explored the role of XAI in education. They have provided a framework with six key aspects in the development of tools used in education. They have analyzed the application of their proposed framework in four AI based educational tools.

Riccordo et.al (2021)have examined the application of AI in socially sensitive and safety-critical contexts. They have introduced a local classifier with identification of neighborhood using genetic algorithm. Then they have generated the factual explanations of the decision taken with a set of counterfactuals which can lead to an alternate outcome.

Frąckiewicz(2023) has examined the application of XAI in education to improve education and personalized learning practices. The emergence of explainable AI (XAI) has enabled educators to create tailored learning experiences for their students. XAI is a type of AI that can explain its decisions and processes, allowing teachers to understand how the AI is making decisions and how it can be used to enhance learning. In addition, XAI can help teachers and students identify patterns and trends in data that may not be immediately obvious. By providing an interactive visual representation of the data, teachers can help students make connections between the data and the underlying concepts. This can help students better understand the material and develop a more comprehensive understanding of the topic.

Farrow (2023) have examined The socio-technical perspective indicates that explicability is a relative term. Consequently, XAIED mediation strategies developed and implemented across education stakeholder communities using language that is not just 'explicable' from an expert or

technical standpoint, but explainable and interpretable to a range of stakeholders including learners. The discussion considers the impact of XAIED on several educational stakeholder types in light of the transparency of algorithms and the approach taken to explanation.

In their research work, Orestis et al.(2021)have aapplied a mode-agnostic approach to elaborate decisions based on the classification of short length words. Their approach generates the text by applying variationalautoencoders to encode and decode the text. Then an AI based learning technique using decision tree is used to generate the neighbors which are then applied to find the exemplars and counter-exemplars.

3. Proposed XAI based Framework

To improve the learning ability and performance improvement of the students, **it is essential to have a learning analytics platform that works with Explainable AI (XAI)**. The proposed solution has the following novel features:

- ➢ a novel solution embedded with **XAI to deliver an intelligent and explainable feedback** to the students to improve their learning ability

- ➢ Development of **ML algorithm led recommendations or actions** by reassuring the stakeholders such as students and parents that there are no biases which may affect the actions to be taken.
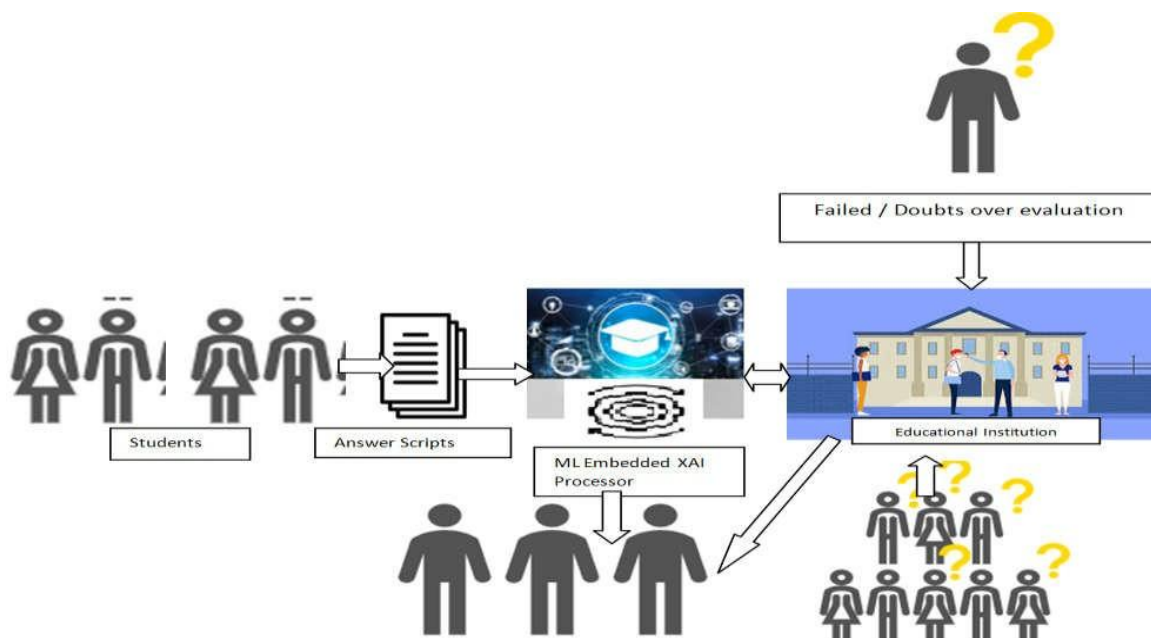
3.1 Proposed Framework

Figure 3 – Proposed XAI based Framework for Assessment in TLP

The framework has the AI engine that helps to dynamically build, manage and follow-up of policies and rubrics to take the decisions. Then a human teacher can easily augment it by providing necessary explanations that are very much appropriate to the corresponding student and the context will also be an expressive one to achieve effective outcome.

This helps in providing the (i) Explainable feedback for the students (ii) Actionable Feedback for the teachers to take decisions (iii) Guided feedback to follow the rubrics insisted.

Hence, the XAI here acts as the bridge between the teacher and the student in order to help improve the students' performance which is the ultimate aim of any effective teaching and learning process.

3.2 Proposed Algorithm for XAI in TLP

Step 1 – **Develop an AI engine** that helps to dynamically build, manage and follow-up of policies and rubrics to take the decisions.

Step 2 – **Provide an UI** through which the human teacher can easily augment the policies and rubrics by providing necessary explanations

Step 3 – Assist the students to enter into the **ML assisted XAI processor** to get personalized explanations associated with the particular student

Step 4 – Showcase possible **improvement guidelines** through the XAI processor to the corresponding student

Step 5 - **Create a context** to be expressive to achieve effective outcome.

4. Case Study

Consider that a student fails in an exam as the corresponding course teacher has refused to give marks for some of the questions. To arrive at this marks, the teacher has a number of data points in terms of rubrics.

But to make the student understands what is wrong and to perform well in the subsequent exams, a contrastive explanation method has to be used. This method will show why marks have been refused and it can also highlight the difference between a successful answering and an unsuccessful answering.

This contrastive explainable method used in XAI helps with the cognitive process of understanding and explaining the results of the problem. The student can then focus on which factors need to be improved to reduce the contrast so that in the next exam he can secure high marks. Simple marking of correct or wrong in the scripts will never help them to improve their standard.

This helps in providing the

- **Explainable feedback** for the students

- **Actionable Feedback** for the teachers to take decisions

- **Guided feedback** to follow the rubrics insisted.

5. Conclusion and Future Work

XAI is the most effective practice to guarantee that AI and ML solutions are transparent. This is trustworthy, responsible, and ethical so that all regulatory requirements on algorithmic transparency, risk mitigation, and a fallback plan are addressed efficiently.AI and ML explainability techniques provide visibility into how algorithms are operated at different stages.XAI is allowing end-users to ask questions about the consequence of AI and ML models.Hence, the XAI here acts as the bridge between the teacher and the student in order to help improve the students' performance which is the ultimate aim of any effective teaching and learning process.As we are applying data-driven XAI, if the data is inconsistent, there are huge chances of the failure of the ML model. The advantages of XAI in TLP include Reduced errors in decision making process, Reduced impact of model bias, Improved model performance, and Guided decision making.

Challenges and Future work

As we are applying data-driven XAI, if the data is inconsistent, there are huge chances of the failure of the ML model. The solution is to apply noise removal techniques to keep relevant data and authorization based real time data to be taken for decision making. In the future work, it is going to be applied in a real time dataset with XAI based implementation.

References

1. Saeed W., Omlin C. (2023), Explainable AI (XAI): A systematic meta-survey of current challenges and future opportunities, Knowledge-Based Systems, Volume 263, 2023, 110273, ISSN 0950-7051, https://doi.org/10.1016/j.knosys.2023.110273

2. Haque B., Islam N., Mikalef P. (2023), Explainable Artificial Intelligence (XAI) from a user perspective: A synthesis of prior literature and problematizing avenues for future research, Technological Forecasting and Social Change, Volume 186, Part A, 2023, 122120, ISSN 0040-1625, https://doi.org/10.1016/j.techfore.2022.122120.

3. Clement T, Kemmerzell N, Abdelaal M, Amberg M (2023). XAIR: A Systematic Metareview of Explainable AI (XAI) Aligned to the Software Development Process. *Machine Learning and Knowledge Extraction*. 2023; 5(1):78-108. https://doi.org/10.3390/make5010006

4. Nwakanma CI, Ahakonye LAC, Njoku JN, Odirichukwu JC, Okolie SA, Uzondu C, NdubuisiNweke CC, Kim D-S. (2023), Explainable Artificial Intelligence (XAI) for Intrusion Detection and Mitigation in Intelligent Connected Vehicles: A Review. *Applied Sciences*. 2023; 13(3):1252. https://doi.org/10.3390/app13031252

5. Riccardo G., Anna M., Fosca G., Dino P., Salvatore R., Franco T. (2021), Factual and Counterfactual Explanations for Black Box Decision Making, IEEE Intelligent Systems. In IEEE Intelligent Systems

6. Theissler A., Spinnato F., Schlegel U., Guidotti R. (2022) , Explainable AI for Time Series Classification: A Review, Taxonomy and Research Directions , IEEE Access. In IEEE Access (Volume: 10)

7. Cecilia P., Andrea B., Daniele F., Fosca G., Dino P., Alan P., Salvatore R. (2022). Co-design of human-centered, explainable AI for clinical decision support In ACM Transactions on Interactive Intelligent Systems

8. Khosravi H., Shum S.B., Chen G., Conati C., Tsai Y., Kay J., Knight S., Martinez-Maldonado R., Sadiq S., Gašević D. (2022), Explainable Artificial Intelligence in education, Computers and Education: Artificial Intelligence, Volume 3, 2022, 100074, ISSN 2666-920X, https://doi.org/10.1016/j.caeai.2022.100074.

9. Farrow R. (2023), The possibilities and limits of XAI in education: a socio-technical perspective, Learning, Media and Technology, DOI: 10.1080/17439884.2023.2185630

10. Frąckiewicz M, (2023), Explainable AI (XAI) in Education: Enhancing Learning and Personalized Instruction, Artificial intelligence, TS2 Space on 6 April 2023

11. Orestis L., Riccardo G., Salvatore R. (2021), Explaining Sentiment Classification with Synthetic Exemplars and Counter-Exemplars, Discovery Science. In In International Conference on Discovery Science (pp. 357-373). Springer, Cham.