Explainable Learning Analytics: challenges and opportunities

Tinne De Laet , Martijn Millecamp, Tom Broos, Robin De Croon, Katrien Verbert
KU Leuven, Belgium
{tinne.delaet,martijn.millecamp,tom.broos,robin.decroon,katrien.verbert}@kuleuven.be

Raphael Duorado
Universidade Federal de Pernambuco, Brasil
rasd2@cin.ufpe.br

ABSTRACT: In the last decade, we are witnessing a widespread adoption of artificial intelligence in a wide range of application domains. Learning analytics is no exception. Artificial Intelligence (AI) techniques and Machine Learning (ML) in particular are used to generate automatic predictions and recommendations regarding learning and teaching. A key challenge in the actual use and adoption of AI and ML is that they often operate as a 'black box', hereby impeding understanding and trust. The domain of Explainable Artificial Intelligence aims at enhancing the transparency of AI techniques and therefore also holds substantial promise for the Learning Analytics domain. This paper supports the shaping of the research line of Explainable Learning Analytics (XLA), by exploring the challenges and opportunities related to the data, stakeholders, communication, evaluation, and implementation & adoption of XLA. In particular, this paper reports on the outcomes of a 3-hour workshop with 44 international participants in which these challenges and opportunities were collaboratively identified. The obtained challenges and opportunities will form the basis for a deeper exploration, involving a wide range of stakeholders, of the promises of the XLA-field and the required points of focus for the next 10 years.

Keywords: learning analytics, explainable learning analytics, explainable artificial intelligence, recommender systems, visual analytics

Figure 1: Example of workshop outcome: opportunities and challenges of explainable learning analytics regarding the stakeholder dimension. The stickers are the result of an up- and downvoting procedure (green = upvoted challenge, gold = upvoted opportunity, yellow = veto).

1 INTRODUCTION

The domain of Learning Analytics (LA) has been established at the background of the never-ending need for support of learners and teachers and under impulse of the growth of learning data, the development of algorithms and AI, and the research within the learning sciences. LA is about "collecting traces that learners leave behind and using those traces to improve learning" (Duval, 2012). Obtaining actual improvement in learning and teaching does not come easy however. Verbert et al. (2013) introduced a LA process model consisting of four levels: awareness (level 1), (self-) reflection (level 2), sensemaking (level 3), and impact (level 4). This model shows that impact in the form of behavioral changes, new meanings, and improvements can only be obtained after awareness of the data, (self-)reflection, and sense-making have been obtained. Machine Learning (ML) and Artificial Intelligence (AI) have provided plenty of algorithms that can support the translation from data to awareness, self-reflection, and sense-making. Visualization techniques in general and Learning Analytics Dashboards (LAD) in particular (Verbert et al., 2014) have been shown to be able to provide a visual means of communication of the data and the outcomes of AI algorithms to stakeholders. While the rapidly maturing LA domain has proven to be an interesting domain of application for AI, ML, and visual analytics, it is creating higher expectations regarding the algorithmic predictions and recommendations generated by AI, ML, and visual analytics. These are expected to be interpretable for and explainable to the involved stakeholders and should be able to be translated to actionable recommendations.

To be **interpretable and explainable**, the outcomes of the data analysis, visualization, and/or ML and AI algorithms often have to be tailored to the particular stakeholders and end-users. While advanced visualization and/or ML techniques might create accurate and trustworthy insights and recommendations, this not automatically leads to the users trusting their outputs. Opening the black-box of LA to the user, in a user-tailored fashion, is an important step towards obtaining interpretable insights and explainable recommendations. The use of new approaches to obtain transparency, trustworthiness, persuasiveness, and effectiveness support this evolution.

A second challenge for LA, after interpretability and explainability, is to translate predictions and recommendations into feasible 'actions'. This is also referred to as **actionability**. To highlight the challenge of this actionability, let's consider the following example. Educational data mining techniques may discover that male students on average are more likely to fail in higher education. While such information can be interesting for researchers and policy makers, it does not provide a directly actionable recommendation towards an individual (male) student on how to improve his learning or study success. If actionable insights and recommendations can be created within LA and if they can be tailored to the involved stakeholders, they will have the potential to create impact (Verbert et al., 2013). User-centered design involving the stakeholders and the integration of LA into actual educational practices and in the pedagogy underlying these educational practices will support the actionability of the insights and recommendations. For example; if instructors collaborate with ML researchers when integrating LA in the form of automatic resource recommendation in their course design, they can help to understand and interpret the automatic recommendation of resources to students in the context of a particular class.

The domain of explainable Artificial Intelligence (XAI) has been developing and growing fast in the last years. AI is a part of a new generation of AI technologies called the third wave AI including, among other ambitious goals, the development of algorithms that can explain themselves (Adadi & Berrada, 2018). The XAI research field aims at improving trust and transparency of AI-based systems, which can both concern automatic predictions and recommendations. AI algorithms often suffer from the so-called 'black-box' phenomenon, indicating that it is hard for users, including domain experts, to get insights in the internal mechanisms underlying the algorithms and the outcomes produced by these algorithms. This is also referred to as algorithm opacity (Adadi & Berrada, 2018). The problem of opacity has been growing together with the development of novel ML algorithms, such as deep learning and random forests, which itself was supported by the rapidly growing computational power. Algorithm opacity can however impede trust in predictions and recommendations provided by these ML algorithms and AI techniques, preventing their actual adoption and deployment in real-world scenarios.

The evolution towards actionable insights and explainable recommendations is urgent, as recent data protection and privacy regulations like the EU General Data Protection Regulation (GDPR) and the California Consumer Privacy Act (CCPA) stipulate that transparency is a fundamental right. In this light, the use of 'black box' approaches towards end-users becomes more and more challenging as the algorithmic approaches themselves lack transparency for LA end-users. Even more, GDPR states that users have the right to withdraw themselves from automatic decision making and profiling. Human mediation in automatic decision making and profiling is a promising approach to accommodate for the ethical use. Human mediation can however only reach its full potential if the algorithmic outputs are interpretable and explainable by the human mediator.

The XAI field by itself is rapidly maturing as shown in the XAI survey of Adadi and Berrada (2018) and by even more recent contributions focusing on the trends within XAI (Abdul et al. , 2018), on the sub-domain of explainable recommendations (Y. Zhang & Chen, 2018; Ouyang, Lawlor, Costa, & Dolog, 2018), the evaluation of XAI (Mohseni, Zarei, & Ragan, 2018), and visual interpretability (Spinner, Schlegel, Schäfer, & El-Assady, 2019; Zhao, Wu, Lee, & Cui, 2019; Q.-s. Zhang & Zhu, 2018; Choo & Liu, 2018).

The survey of Adadi and Berrada (2018) also recognizes the potential for XAI in different application domains: transportation, healthcare/medical (Vellido, 2019; Kwon et al., 2019), legal, finance, and military. The number of application domains is still growing fast as shown by recent research dedicated to e.g. robotic agents (Anjomshoae, Najjar, Calvaresi, & Frmling, 2019). While attention for XAI is also growing within the domain of LA, it still remains to be determined what the main research directions should be, to what level general XAI findings can be applied to the LA domain, and to what level specific developments have to be made.

The goal of this paper is to contribute to the creation and shaping of the exciting and promising domain of Explainable Learning Analytics (XLA), focusing on the application domain-specific developments of XAI within LA. In particular, this paper aims at contributing to the discovery of the main opportunities and challenges of XLA. To this end, this paper reports on a workshop involving more than 40 international stakeholders to identify the main challenges and opportunities of XLA.

2 METHODOLOGY

Stakeholder input regarding the challenges and opportunities of XLA was collected during a 3-hour workshop at the 2019 European Conference on Technology Enhanced Learning (EC-TEL 2019), in Delft, the Netherlands. Beforehand, the involved researchers identified five themes of focus for the workshop: data, stakeholders, communication, evaluation, and implementation & adoption. Table 1 provides an overview of the themes and how they were presented during the workshop.

Table I: The five different themes of focus of explainable learning analytics and the teaser questions that were provided to the participants.

Theme	Explanation verbally provided to participants
Data	What data can be useful to be explained? What data about the user can be used to generate a prediction or a recommendation? Is the data readily available?
Stakeholders	Who are the stakeholders of explainable LA? Why would they use LA that needs explanations? In which situation would they need explanations? When would they see/use these explanations?
Communication	How would you communicate explanations to the user? (visual, text, mix, audio, ...). How would you adapt the explanation? Based on personal/situational characteristics? Is it ethical to adapt interfaces? Why would you trust or not trust a system? Do explanations help?
Evaluation	How would you evaluate explanations? Qualitative? Quantitative? What end goal would you evaluate?
Implementation & adoption	What steps are needed to implement XLA? Where do you see the domain of XLA in 10 years? What is needed to reach that? What can inhibit XLA? What can stimulate XLA? For what purpose would you use XLA?

These themes are inspired by the six critical dimensions of LA of Greller & Drachsler (2012) (data, stakeholders, instruments, internal limitations, external constraints) supplemented with and made more concrete by experience of the researches themselves in the implementation at institutional scale of student dashboards (Broos et al., 2020) . The data and stakeholder theme were directly borrowed from the six critical dimensions. Communication is one specific aspect of the critical dimension 'instruments', focusing on the communication of algorithmic predictions and recommendations in the context of the workshop. The theme of implementation & adoption touches on the internal limitations and external constraints of the six critical dimensions. However, we decided, based on our experience with deploying learning dashboards at institutional scale to focus specifically on implementation & adoption. Finally, the theme of evaluation concerns both the evaluation methodology (Instruments dimension), but also what final objective (Objectives dimension) to evaluate. These themes were used to structure the conversation and we do not claim that these five themes entail all possible viewpoints of XLA, nor that they are the only way to structure them.

During the workshop the following protocol was used: (1) Welcome and ice-breaker activity (10 minutes); (2) Introduction regarding LA, explainabilty and interpretability of predictions and recommendations (15 minutes), (3) Idea generation round in small groups (60 minutes, Figure 2), (4) Synthesis round where all input per theme was collected and prioritized using grouping of input and up and downvoting, (30 minutes, Figure 1), (5) Plenary discussion to finalize the identified challenges and opportunities (65 minutes).

Figure 2: Picture of the idea generation round during the workshop at the EC-TEL 2019 conference

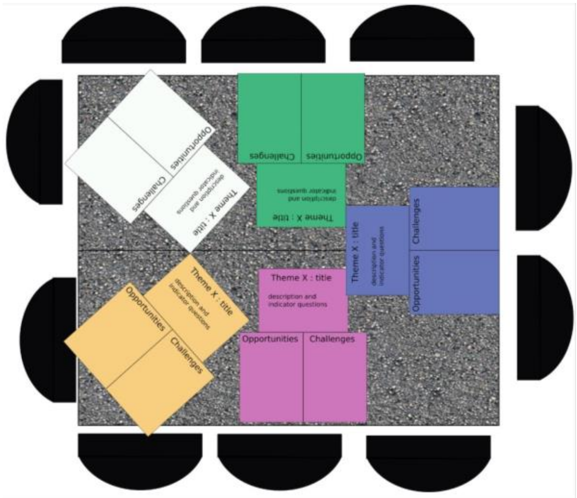

Figure 3: Table lay-out for the idea generation round. At each table the five themes are discussed by pairs of participants.

In the *introduction* a plenary presentation was provided with the general background of explainable AI, LA, and some examples of XLA. The presentation also focused on the concepts of predictions, recommendations, explainability, and interpretability and why these are important. This introduction ensured that each participant had a basic understanding of the field of XAI and LA. Additionally, everyone was made aware of the protocol used in the workshop.

The *idea generation round* was organized such that pairs of participants discussed around each of the five themes. Tables were prepared to support this, as illustrated in Figure 3. Pairs of participants would discuss on the challenges and opportunities related to a particular theme during around twelve minutes using a push-through procedure. They added the output of their discussion using post-its to the discussion notes on the table. Each twelve minutes, the pairs progressed to the next theme at the same table and would add their findings to the already existing discussion notes. This round ended as soon as each pair of participants had addressed each theme once.

In the *synthesis round* the input from the different tables was grouped thematically, i.e. according to each of the five themes. First, the participants were asked to group the input (post-its) by grouping similar ideas. Next, participants were invited to individual dot-voting: each participant received eight stickers that he/she could use to highlight the most urgent or important challenges (4 stickers) and opportunities (four stickers). Additionally, each participant received two veto stickers to indicate their disagreement: one for an opportunity and one for a challenge. Participants were requested to put their initials on the veto sticker such that they could be prompted for more explanation during the plenary discussion round.

The goal of the *plenary discussion* was to use the input from the idea generation and synthesis rounds to define the most important challenges and opportunities of XLA for each of the five themes. A list of most important challenges and opportunities was assembled based on the number of dot votes. Selected items for each of the themes were discussed one by one, and participants were invited to bring forward the findings and elaborate on them. Other participants were invited to comment and discuss. This discussion was recorded using audio recording devices.

Ethical approval was obtained from the ethical commission of KU Leuven. Workshop participants that subscribed to the workshop received an explanation of the research performed beforehand by email, including a notice that they would be invited to sign a consent form if they would participate in the plenary discussion. During the workshop, all participants were informed about the protocol and invited to sign the consent forms if they agreed their audio was recording during the plenary discussion. Participants not consenting could still participate in all parts of the workshop, except the final audio-recorded plenary discussion. At several stages during the workshop pictures were taken to support the processing of results. Prior consents was requested for participants being pictured.

3 RESULTS

The workshop organized at the EC-TEL 2019 conference in Delft, the Netherlands was attended by 44 participants. All attendants participated in the idea generation round. After a short break, 17 participants consented with the recording of their voices during the final discussion, and therefore participated in both the idea generation round, the synthesis round, and the plenary discussion. Below, we elaborate on the main opportunities and challenges that were identified during the synthesis round and the plenary discussion (as illustrated in Table 2), grouped by each of the five themes.

Table II: Most voted challenges and opportunities.

Theme	Item	#opportunity	#challenge	#veto
	Opportunities			
Data	Course/learning design, re-using teachers' previous data	6	0	0
Stakeholders	Teachers reflect/understand own effectiveness (by visualization features)	4	1	0
Communication	Data storytelling	7	0	0
Evaluation	Evaluate impact of the explanations	3	1	0
Implementation & adoption	Adapt to the target groups	5	2	0
	Challenges			
Data	Include how recent the data is	0	4	0
Stakeholders	Community building	1	4	0
Communication	Time-based LA, splitting explanation per phase/step	0	6	0
Evaluation	Added value for user (pre/post)?	1	3	0
Implementation & adoption	Support (technical, pedagogical) when system is deployed	1	3	0

3.1 Opportunity

Data. The main identified opportunity for data for XLA is to use data regarding course/learning design and reusing teachers' previous data (6 opportunity stickers, no challenge stickers, no veto).

During the discussion, participants elaborated that if humans generate the data, this might alleviate the explainability issue. After all, human-generated data could be useful to explain computer-generated data. The participants are interested to know how learning design, if well-modelled, will influence predictions and explanations and how feedback regarding these predictions and explanations can improve the underlying models. An example of such an improvement is elaborated on in the paper of Mothilal et al. (2018), where the explanatory technique of LIME is used to obtain explanations of the prediction of first-year engineering student success, which improved the model of student success. It is however still an open research challenge what data can and should be collected to support explainability. Nonetheless, the participants agreed that the expectations should be clear beforehand and that annotations on the data, if used with care, can support the explainability. A final warning regarding the data of course/learning design is that it can and should be context-specific, complicating wider use.

Stakeholders. The main identified opportunity of XLA for stakeholders within the workshop is for teachers when they can use XLA to reflect upon or understand their own effectiveness. (4 opportunity stickers, 1 challenge sticker, no veto). XLA will definitely provide an opportunity for teachers as XLA can disclose understandable explanations and recommendations to teachers regarding their teaching effectiveness. Participants also remarked that it would be a challenge however combine different perspectives of stakeholders, especially if they are conflicting. How should the perspectives be prioritized or weighted?

Communication. The main opportunity for the communication of XLA is data storytelling (7 opportunity stickers, 0 challenge stickers, no veto). Recent research and technological advancement have identified opportunities to automate data storytelling (Echeverria et al. 2018). It remains to be researched, however, to what level this automation is feasible and how and when a data scientist should still be in the loop. Storytelling, both manual and automatic, unlocks the opportunity of personalization. This immediately raises additional concerns related to ethics. For example, can personalized explanations trigger different interpretations depending on the personalization? Storytelling also has the opportunity to emphasize particular parts of the data, hereby providing an answer to the data abundance problem. One should be careful, however, not to 'obscure' the data: one should be transparent on which data is emphasized and which is hidden.

Evaluation. The main identified opportunity of the evaluation of XLA is related to the evaluation of the impact of the explanations (3 opportunity stickers, 1 challenge sticker, 0 veto). While the problem of evaluating XLA is new, there are consolidated techniques that could be adapted. Before evaluation can be started however, it is important to define clearly the different evaluation goals: they can range from perceived utility to impact on, e.g., advising and decision making. One should take care to not only set up separate evaluations with the different stakeholder groups, but to use the opportunity of mixed-group evaluations. The evaluation should moreover focus on both subjective and more objective indicators: one should not only rely on subjective statements but also attempt to look for objective/quantitative measures, such as the impact on learning gain.

Implementation & adoption. The main opportunity for the implementation and adoption of XLA is the adaptation to different target groups (5 opportunity stickers, 2 challenge stickers, no veto) Different target groups might need different types of explanations and interpretations of LA predictions and recommendations. Each stakeholder might have particular needs and therefore, the explanations and interpretations should be personalized to the particular group of stakeholders.

3.2 Challenges

Data. The main identified challenge for data in XLA is to include information about and take into account how recent the data is (0 opportunity stickers, 4 challenge stickers, 0 veto stickers). It is challenging to find a threshold that could uniquely define what 'recent' and 'old' data are. There is a tension between how valuable old data (to obtain enough data to train the models or to show historical evolution) and new data (more representative of current state) is. Additionally, attention should be paid to how to explain to users which data is used in these models and that predictions and recommendations rely on past data. Finally, deploying XLA can, and most likely will, influence the data itself as it is expected to have an impact on actual learning and teaching.

Stakeholders. The main challenge for the stakeholders is to build a strong community. (1 opportunity sticker, 4 challenge stickers, no veto). For the entire field of LA it is a challenge to build a strong community that could support stakeholders working on explanations of predictions and recommendations en strengthen their collaboration and the adoption of XLA When done well, the explanations have the opportunity to foster trust among different stakeholders, for example among students and teachers in a MOOC.

Communication. The main challenge for communication within XLA is to consider the time dimension of learning analytics (0 opportunity stickers, 6 challenge stickers, 0 veto). Longitudinal data is challenging to handle within learning analytics. XLA should be able to provide explanations for the different phases over time. Moreover, these explanations should be tailored to the particular phases and contexts they are provided in.

Evaluation. The main challenge for a good evaluation of XLA is to identify the added value for the stakeholders (1 opportunity sticker, 3 challenge stickers, 0 veto). The evaluation of XLA should focus on identifying the added value of explanations for different stakeholders, and in particular should be able to show how the explanations contribute to what the stakeholders already know (e.g., using a pre/post test setup).

Implementation & adoption. The main challenge for the implementation and adoption of XLA are both technical and pedagogical support during deployment (1 opportunity sticker, 3 challenge stickers, 0 veto). The actual implementation and adoption of XLA will provide ample challenges, especially when deployments at scale are considered. These issues are not only of technical nature, but also pedagogical: how can the explanations be used appropriately during the learning process?

4 DISCUSSION AND CONCLUSION

This short paper called upon the input of more than 40 stakeholders to shape the domain of Explainable Learning Analytics (XLA), which aims at developing LA-specific advancement regarding Explainable Artificial Intelligence (XAI). In particular, this paper reports on the opportunities and challenges of XLA as identified by this group of international stakeholders collected during a 3-hour workshop at the 2019 European Conference on Technology Enhanced Learning (EC-TEL). Opportunities and challenges were collected regarding five main themes: data, stakeholders, communication, evaluation, and implementation & adoption. The next step in future research would be to make a deeper analysis of the workshop outcomes and especially the audio recordings made and then to compare the outcomes to existent findings in XLA and XAI, for instance to the findings of Miller (2017) and Karga & Satratzemi (2019).

The input from stakeholders is undoubtedly valuable for the advancement of XLA. This contribution is, due to several limitations, only a small step towards a more profound integration of the different stakeholders in the development of the domain. A first limitations is that the workshop was held at the EC-TEL 2019 conference hereby causing a biased sample of the stakeholder population. The involved stakeholders were mainly researchers active in Technology Enhanced Learning, possibly causing the observed bias towards teachers in for instance the stakeholders dimension. Future stakeholder consultations should more heavily involve practitioners, policy-makers, and end-users of LA. Earlier work will provide inspiration regarding for instance the inclusion of students as a stakeholder in XLA (Putham & Conati 2019; Baria-Pineda & Brusilovsky 2019). A positive element of the stakeholder population is that they represented the wider domain of Technology Enhanced Learning, of which LA is only a sub-domain. On the negative side however, hereby introducing the second limitation, this meant that some attendants were not very acquainted with the specifics of the LA domain, while others were considered experts. The same holds for XAI: some attendants were experienced researchers or users, while others were not familiar with the domain. For future stakeholder consultations we recommend to set up a protocol that aims at better handling such differences in expertise, both regarding the LA and the XAI domain. A third limitation is the short duration of the workshop, which limited both the width and the depth of the discussion, and the limitations of the recordings made (only during the synthesis round).

To conclude, we can state that this paper contributes to the development of the XLA domain by the identification of challenges and opportunities regarding data, stakeholders, evaluation, communication, and implementation & adoption.

ACKNOWLEDGMENTS

We thank the participants of the EC-TEL 2019 XLA workshop for their contribution. This work was funded by the EU LALA project (grant no. 586120-EPP-1-2017-1-ES-EPPKA2-CBHE-JP). This project has been funded with support from the European Commission. This publication reflects the views only of the authors, and the Commission cannot be held responsible for any use which may be made of the information contained therein.

REFERENCES

Abdul, A., Vermeulen, J., Wang, D., Lim, B. Y., & Kankanhalli, M. (2018). Trends and trajectories for explainable, accountable and intelligible systems: An hci research agenda. In Proceedings of the 2018 chi conference on human factors in computing systems (pp. 582:1–582:18). New York, NY, USA: ACM. Retrieved from http://doi.acm.org/10.1145/3173574.3174156 doi: 10.1145/3173574.3174156

Adadi, A., & Berrada, M. (2018). Peeking inside the black-box: A survey on explainable artificial intelligence (xai). IEEE Access, 6, 52138–52160.

Anjomshoae, S., Najjar, A., Calvaresi, D., & Frmling, K. (2019). Explainable agents and robots: Results from a systematic literature review. In (pp. 1078–1088).

Barria-Pineda, J. & Brusilovsky, P. (2019) Making Educational Recommendations Transparent through a Fine-Grained Open Learner Model. IUI Workshops 2019

Broos, T., Pinxten, M., Delporte, M., Verbert, K., & De Laet, T. (2019). Learning dashboards at scale: early warning and overall first year experience. Assessment & Evaluation in Higher Education, 1–20. https://doi.org/10.1080/02602938.2019.1689546

Choo, J., & Liu, S. (2018). Visual analytics for explainable deep learning.

Duval Erik (2012). Learning Analytics and Educational Data Mining, Erik Duval's Weblog, 30 January 2012, https://erikduval.wordpress.com/2012/01/30/learning-analytics-and-educational-data-mining/

Echeverria, Vanessa and Martinez-Maldonado, Roberto and Granda, Roger and Chiluiza, Katherine and Conati, Cristina and Shum, Simon Buckingham (2018). Driving Data Storytelling from Learning Design. Proceedings of the 8th International Conference on Learning Analytics and Knowledge (LAK 2018) Sydney, Australia, isbn: 978-1-4503-6400-3}, (pp 131—140), doi: 10.1145/3170358.3170380

Greller, Wolfgang & Drachsler, Hendrik. (2012). Translating Learning into Numbers: A Generic Framework for Learning Analytics. Educational Technology & Society. 15. 42-57.

Karga, S., & Satratzemi, M. (2019). Using explanations for recommender systems in learning design settings to enhance teachers' acceptance and perceived experience. Education and Information Technologies, 24(5), 2953–2974. https://doi.org/10.1007s10639-019-09909-z

Kwon, B. C., Choi, M.-J., Kim, J. T., Choi, E., Kim, Y. B., Kwon, S., ... Choo, J. (2019). Retainvis: Visual analytics with interpretable and interactive recurrent neural networks on electronic medical records. IEEE Transactions on Visualization and Computer Graphics, 25(1), 299–309.

Miller, T. (2017). Explanation in Artificial Intelligence: Insights from the Social Sciences. Artificial Intelligence, 267, 1–38. https://doi.org/arXiv:1706.07269v1

Mohseni, S., Zarei, N., & Ragan, E. (2018). A survey of evaluation methods and measures for interpretable machine learning. arXiv.org. Retrieved from http://search.proquest.com/docview/2139809046/

Mothilal, R. K., & Broos, T. (2018). Predicting first-year engineering student success: from traditional statistics to machine learning. In Proceedings of the 46th sefi conference (pp. 1–8). Copenhagen, Denmark. Retrieved from https://www.sefi.be/proceedings/?conference=copenhagen2018

Ouyang, S., Lawlor, A., Costa, F., & Dolog, P. (2018). Improving explainable recommendations with synthetic reviews.

Putnam, V., & Conati, C. (2019). Exploring the Need for Explainable Artificial Intelligence (XAI) in Intelligent Tutoring Systems (ITS). IUI Workshops 2019

Spinner, T., Schlegel, U., Schäfer, H., & El-Assady, M. (2019). explainer: A visual analytics framework for interactive and explainable machine learning. IEEE transactions on visualization and computer graphics.

Vellido, A. (2019). The importance of interpretability and visualization in machine learning for applications in medicine and health care. Neural Computing and Applications.

Verbert, K., Duval, e., Klerkx, J., Govaerts, S., & Santos, J. (2013, 09). Learning analytics dashboard applications. American Behavioral Scientist, 57. doi: 10.1177/0002764213479363

Verbert, K., Govaerts, S., Duval, E., Santos, J. L., Van Assche, F., Parra, G., & Klerkx, J. (2014). Learning dashboards: an overview and future research opportunities. Personal and Ubiquitous Computing, 18(6), 1499–1514.

Zhang, Q.-s., & Zhu, S.-c. (2018). Visual interpretability for deep learning: a survey. Frontiers of Information Technology & Electronic Engineering, 19(1), 27–39.

Zhang, Y., & Chen, X. (2018). Explainable recommendation: A survey and new perspectives.

Zhao, X., Wu, Y., Lee, D. L., & Cui, W. (2019). iforest: Interpreting random forests via visual analytics. IEEE Transactions on Visualization and Computer Graphics, 25(1), 407–416.