Workshop on Learning Analytics

Considering student diversity with regard to assessment data and discrimination.

Clara Schumacher[1], Nathalie Rzepka[2] und Niels Seidel[3]

Abstract: This workshop focuses on current topics in the domain of learning analytics (LA). In particular, the increasing diversity of students in higher education also needs to be considered in LA. Offering adaptive support is a major aim of LA. Assessments are considered having a major impact on student learning. However, to date LA still do not sufficiently make use of assessment data related to context, activity and results. In contrary to the aim of giving the individual learner the support needed, LA are also criticized to foster prejudices. This becomes even more important in the light of a high diversity of the student cohorts. Thus, this workshop includes submissions in the domain of LA with emphasize on the impact of outliers on dropout prediction, students' perceptions of algorithms regarding grading, students' control over data collection, the application of the FAIR principles for data management to learning analytics, and the identification of indicators of group learning in collaborative software development. Furthermore, Professor Ryan Baker from University of Pennsylvania will hold a keynote on "Algorithmic Bias in Education".

Keywords: learning analytics, diversity, assessment analytics, discrimination

1 Introduction

Due to higher enrollment rates, student diversity is gaining relevance in higher education [HMU17]. The term diversity in higher education is used in the context of underrepresented student groups [Bo15] but refers also to students' individual differences regarding for instance prior education, learning strategies and motivation, and preferences [ZM14].

To offer increasingly diverse students the support needed, LA might be a suitable means as they aim at offering feedback and identifying students at-risk for timely interventions. However, to offer valid feedback on learning processes and on how to improve data collection and analysis need to be guided by theory [Se20]. Assessments are considered to determine what students are learning. Hence, assessments should not focus on tasks that are easy to analyze such as single- or multiple-choice but on tasks assessing valued learning outcomes [SB10]. Considering the debate on 21st century skills the valued

[1] Humboldt-Universität zu Berlin, Department of Computer Science, Unter den Linden 6, Berlin, 10099, clara.schumacher@hu-berlin.de
[2] Hochschule für Technik und Wirtschaft Berlin, Treskowallee, 8, 10318 Berlin, nathalie.rzepka@htw-berlin.de
[3] FernUniversität in Hagen, Chair of Cooperative Systems, Universitätsstraße 1, 58084 Hagen, niels.seidel@fernuni-hagen.de

competences are complex to measure [We18]. In that respect, [Lo17] emphasize the need for further interdisciplinary efforts for integrating assessment models and research with LA. As a first attempt, [NCB19] developed an ontological model enhancing current xAPI approaches with contextual assessment data (e.g., assessment type; prior knowledge; current assessment performance; difficulty level; assessment results and attempts). Only when the data collected are valid, interpreted with relation to theory, and their context meaningful inferences and interventions can be deduced from LA. The validity is increasingly relevant as LA are used to identify students at-risk which might negatively affect their motivation and self-efficacy beliefs. Therefore, the validation of the analyses must consider potential discrimination of the increasingly diverse students. Neglecting the potential of discrimination can reinforce it by confronting individuals with inappropriate interventions [AP20]. However, just as serious as an inappropriate intervention can be a lack of intervention, whereby a student does not receive the adjustments needed to increase learning success.

For at-risk predictions, the technology applied as well as the variables used are important [SHR15]. Particularly, with regard to discrimination it is vital which variables were used for the algorithms. Most significant are students' GPA and internal assessment, but also demographic data (e.g., age, gender, disability, family background) are used commonly [SHR15]. [GBB19] found significant differences of model fairness for gender which depends on both, the algorithm as well as the feature set used. [RSS20] compared data of disabled and non-disabled students and found, that underrepresentation of minorities can lead to unfair classification. Further research indicated that not only unfair classification can be discriminatory, but also that the quality of predictive models can vary among sub-groups [RS19].

LA are far from being perfect and the context and processes they focus on are complex [KSG18]. Thus, the stakeholders facing LA and the related interventions need to be capable of understanding the underlying basic concepts, the possibilities but also the associated limitations and biases.

Current research might investigate how available assessment models need to be adapted to be integrated into holistic LA and what this means for the learning design of courses. How do current approaches deal with incomplete data in their models and potential self-selection biases of students agreeing to divulge personal data? How are biases of algorithms and discrimination considered in current LA systems in different countries? To what extent are the stakeholders of LA aware of biases and how can this be fostered? Is diversity considered in LA interventions without fostering discrimination but still supporting individual needs? Which variables are necessary to model student diversity without discrimination in LA? How can interdisciplinary communication and research be fostered?

2 Goals of the workshop

The overall goals of this interdisciplinary workshop are: (i) Networking of the LA community in the German speaking countries to initiate research and joint projects as well as fostering network activities beyond this workshop; (ii) Presentation of current research projects and results; (iii) Enhancing the interdisciplinarity in the working group: e.g., computer sciences, artificial intelligence, math, learning sciences and psychology as well as ethics and philosophy; (vi) Developing a joint outcome (e.g., paper, digital learning resources, event) based on the workshop discussions.

3 Submissions to the workshop

The call for papers for the workshop has been published online: http://akla.f4.htw-berlin.de/workshop-2021/ and allowed submissions presenting work in progress and preliminary results to the community. Submissions have been peer-reviewed. In total, 6 submissions have been accepted for publication. The accepted papers deal with topics related to control over data collection, influence of data quality and completeness on prediction, students' perceptions of learning analytics and data collection as well as analysis of students' collaborative software development.

A predominant aim of learning analytics is dropout prediction. However, this might be affected by outliers of student characteristics. Hence, Daria Novoseltseva, Kerstin Wagner, Agathe Merceron, Petra Sauer, Nadine Jessel, and Florence Sedes investigate what kinds of student-outliers can be detected and if they affect dropout prediction by analysing different models using three types of algorithms. Their results indicate that removing outliers could improve dropout prediction.

Grading is work intensive and learning analytics might be a supportive means. In their study Linda Mai, Alina Köchling, Lynn Schmodde and Marius Wehner investigate students' perceptions of three scenarios (a) teacher, (b) teacher supported by learning analytics and (c) learning analytics grading. Their results indicate that students' stated satisfaction is lower with the decrease of teacher involvement in the grading process, but the perception of fairness is higher when grading is based on an algorithm. Furthermore, they found that perceived fairness is a mediator for satisfaction when grading is based on an algorithm.

In the context of adult education Philipp Krieter, Michael Viertel and Andreas Breiter used interviews, log file data and screen recordings to investigate how student privacy can be supported by giving students control over their data collection. Students used tablets for learning music and were allowed to turn on and off the permanent recording of their screen but not the additionally collected system log files. Findings of this small-scale study indicate that this approach does not affect the usefulness of the data set and that students consider it meaningful for preventing the collection of activities not related to learning.

As learning analytics are relying on a great variety of different types of data but common practices for data handling are missing, Ian Wolff, David Broneske and Veit Köppen relate the FAIR principles for data management to the domain of learning analytics.

In their paper Benjamin Weiher, Niels Seidel, Marc Burchart and Dirk Veiel focus on indicators for identifying group learning in collaborative software development. They found a total of 31 indicators for group learning: 7 indicators for code quality, 16 indicators for participation in the group work and 8 indicators on group cohesion. These indicators can be made available on dashboards for supporting teachers in keeping track of group processes and students' individual contributions as well as using the indicators for the provision of feedback.

As programming in teams is considered essential for professional programmers this is also part of higher education courses. However, students' participation and performance in group projects is difficult to measure. Thus, Maximilian Karl and Niels Pinkwart use GitHub data from programming projects to classify students' type of teamwork (e.g., collaboration, cooperation or solo). Their classification method revealed that in their sample most projects were done by one individual student. Such additional information on student contribution can be made available for supporting grading or offering feedback.

The keynote of Professor Ryan Baker, Associate Professor at University of Pennsylvania and Director of the Penn Center for Learning Analytics, is about "Algorithmic Bias in Education". As announced by Ryan Baker his keynote will include the following topics: "The advanced algorithms of learning analytics and educational data mining underpin modern adaptive learning technologies, for assessment and supporting learning. However, insufficient research has gone into validating whether these algorithms are biased against historically underrepresented learners. In this talk, I briefly discuss the literature on algorithmic bias in education, reviewing the evidence for how algorithmic bias impacts specific groups of learners, and the gaps in that literature – both in terms of „known unknowns" and „unknown unknowns". I conclude with potential directions to move the research community towards better understanding how bias impacts educational algorithms, and how to address these problems so that learning systems better promote fairness and equity."

4 Workshop summary

The workshop took place as an all-day online event with up to 25 participants. In the morning the workshop organizers presented the schedule and introduced the topic of the workshop followed by the first paper presentations and their discussion. In order to stimulate further debates, a group phase with a subsequent discussion format using shared editors and breakout rooms guided by statements on current issues of learning analytics took place. After that, the remaining papers were presented and discussed. During the lunch break, there was the opportunity to follow the short paper presentations. In the afternoon, Professor Ryan Baker (University of Pennsylvania) hold a keynote on “Algorithmic Bias in Education”. A stimulating Q&A with Ryan Baker was followed by a summing up discussion with all participants. Furthermore, joint projects and the organization of the working group were discussed.

During the workshop, many discussions and starting points for further research developed. It became clear in many discussions how much impact the rationale for collecting data for LA has on users. If collected data are used for anonymized analysis, this has a very different impact on perceptions of privacy and fairness than if the data are also used to assign grades. In the debate about fairness and satisfaction around grading, there was also an appeal to see greater opportunities in LA applications. Even the subjective grading of teachers is not without its faults, but it seems that researchers in the LA field nevertheless have a particularly cautious view of LA applications, even though they can achieve a lot.

In the group phase, it was discussed to what extent relevant indicators can be defined and measured and how these measurement models are developed. Here it becomes apparent that the application area of LA is so broad and diverse that such questions are difficult to answer globally. It is therefore important to look at the exact task in order to be able to create applications well. Furthermore, indicators should be defined from day-to-day practice and continuously evaluated.

Furthermore, two paper presentations considered collaboration in software development teams presenting different measures for monitoring group activity and progress using data from GitLab and GitHub. In order to provide formative feedback to the learner and to achieve an equal participation in the learning process teachers require support for a wide range of didactic group scenarios and collaborative programming tasks. An open question is the standardization of data formats and sharing of data sets to study group collaboration in small and large software development teams.

Ryan Baker showed in his keynote that fairness and equity can only be driven if one also knows which subgroups are affected by discrimination. He presented examples from literature in which models worked less well for subgroups than for groups with whose data they were trained or tested. He said that there has been no research on algorithmic bias using data from European educational institutions. Apart from this, the research on discrimination in LA mainly focuses on gender and ethnic discrimination, but there are many groups that have not been studied at all. Here he gives the examples of non-binary people, transgender people, international students, language dialects, different educational backgrounds of parents or people with disabilities. In order to move from this unknown bias to known bias, he suggests tackling two main obstacles: lack of data on group membership and the lack of transparency on bias and group-specific outcomes. The lack of data on group membership often results from privacy concerns and legal restrictions. While the students’ privacy should be well taken care of, Baker suggests balancing the right to privacy by the right not to be discriminated against. He concludes that much more data and information should be available on the students’ demographic in order to avoid unconscious discrimination. The subsequent discussion showed here that the transferability of his idea in Europe is only partly realistic, as the differences in the education system and in the handling of personal data are great. The second obstacle, lack of transparency on bias and group-specific outcome should be tackled by defining standards for demonstrating effectiveness, for example in algorithmic bias reviews.

At the end of the event, the importance of interdisciplinarity was highlighted and discussed. In addition, the great responsibility that comes with having an influence on (young) people in the field of education was emphasized. Learning analytics should not be an end in itself, but should support and promote the development of students. This is realized through interdisciplinary projects, which are composed of both the technical as well as the educational practice and pedagogical perspectives.

5 Organizing team (OT) and program committee (PC)

Dr. Clara Schumacher, Humboldt-Universität zu Berlin (OT)
Nathalie Rzepka, Hochschule für Technik und Wirtschaft Berlin (OT)
Dr. Niels Seidel, FernUniversität in Hagen (OT)
Prof. Dr. Niels Pinkwart, Humboldt-Universität zu Berlin (PC)
Prof. Dr. Dirk Ifenthaler, Universität Mannheim (PC)
Prof. Dr. Katharina Simbeck, Hochschule für Technik und Wirtschaft Berlin (PC)
Prof. Dr. Josef Guggemos, Universität St. Gallen (PC)
Dr. Jakub Kuzilek, Humboldt-Universität zu Berlin (PC)
Dr. Martin Hlosta, The Open University UK (PC)
Maximilian Karl, Humboldt-Universität zu Berlin (PC)

Bibliography

[AP20] Archer, E.; Prinsloo, P.: Speaking the unspoken in learning analytics: troubling the defaults. In Assessment & Evaluation in Higher Education, 2020, 45; pp. 888–900.

[Bo15] Bosse, E.: Exploring the role of student diversity for the first-year experience. In Zeitschrift für Hochschulentwicklung, 2015, 10.

[GBB19] Gardner, J.; Brooks, C.; Baker, R.: Evaluating the Fairness of Predictive Student Models Through Slicing Analysis: Proceedings of the 9th International Conference on Learning Analytics and Knowledge. ACM, New York, NY, USA, 2019.

[HMU17] Holmegaard, H. T.; Madsen, L. M.; Ulriksen, L.: Why should European higher education care about the retention of non-traditional students? In European Educational Research Journal, 2017, 16; pp. 3–11.

[KSG18] Kitto, K.; Shum, S. B.; Gibson, A.: Embracing imperfection in learning analytics: Proceedings of the 8th International Conference on Learning Analytics and Knowledge. ACM, 2018.

[Lo17] Lodge, J. M. et al.: Inferring Learning from Big Data: The Importance of a Transdisciplinary and Multidimensional Approach. In Technology, Knowledge and Learning, 2017, 22; pp. 385–400.

[NCB19] Nouira, A.; Cheniti-Belcadhi, L.; Braham, R.: An ontology-based framework of assessment analytics for massive learning. In Computer Applications in Engineering Education, 2019, 27; pp. 1343–1360.

[RS19] Riazy, S.; Simbeck, K.: Predictive Algorithms in Learning Analytics and their Fairness. In 1617-5468, 2019.

[RSS20] Riazy, S.; Simbeck, K.; Schreck, V.: Fairness in Learning Analytics: Student At-risk Prediction in Virtual Learning Environments: Proceedings of the 12th International Conference on Computer Supported Education. SCITEPRESS - Science and Technology Publications, 2020.

[SB10] Shute, V. J.; Becker, B. J.: Prelude: Assessment for the 21st Century: Innovative Assessment for the 21st Century. Springer, Boston, MA, 2010; pp. 1–11.

[Se20] Sedrakyan, G. et al.: Linking learning behavior analytics and learning science concepts: Designing a learning analytics dashboard for feedback to support learning regulation. In Computers in Human Behavior, 2020, 107; 105512.

[SHR15] Shahiri, A. M.; Husain, W.; Rashid, N. A.: A Review on Predicting Student's Performance Using Data Mining Techniques. In Procedia Computer Science, 2015, 72; pp. 414–422.

[We18] Webb, M. E. et al.: Challenges for IT-Enabled Formative Assessment of Complex 21st Century Skills. In Technology, Knowledge and Learning, 2018, 23; pp. 441–456.

[ZM14] Zervakis, P.; Mooraj, M.: Der Umgang mit studentischer Heterogenität in Studium und Lehre. Chancen, Herausforderungen, Strategien und gelungene Praxisansätze aus den Hochschulen. In Zeitschrift für Inklusion, 2014.