Teaching Data Analytics in Higher Education:
The Benefits and Pitfalls of Learning Analytics

Reijonen Atte, South-Eastern Finland University of Applied Sciences, Finland
Helander Nina, South-Eastern Finland University of Applied Sciences, Finland
Hurri Päivi, South-Eastern Finland University of Applied Sciences, Finland
Ansamäki Jarkko, South-Eastern Finland University of Applied Sciences, Finland

The Barcelona Conference on Education 2024
Official Conference Proceedings

Abstract
Modern organizations build their competitive advantage on digital technologies and possibilities provided by data. To succeed in this kind of modern work life, students need various digital skills, including contemporary data analytics skills. Especially there is lack of innovative learning environments and different kinds of hands-on experiences that would transform the teaching of statistics and data analytics towards more applied and critical approaches, favoring the humanistic thinking of students. Learning analytics data can be used to enhance the learning of the students and to produce automatic feedback and adaptive tasks to guide the students in the right direction. With the help of data, the teacher deepens the understanding of the learner's experience and develops teaching. In this paper, the challenges of teaching technical content, such as data analytics, in online course environment are studied. The special interest is to study how learning analytics can be used in teaching technical content for student for pedagogical development. The aim is to provide understanding both on the benefits and the pitfalls, and to develop potential solutions to overcome the challenges. The analysis of the usefulness of learning analytics is carried out from the teachers' perspective. Learning analytics experiences from two online data analytics courses are gathered in the form of teacher interviews. In the end of the paper, guidelines for developing data sources for learning analytics for online environments are provided.

Keywords: Learning Analytics, Data Sources, Data Analytics Teaching, Online Teaching

Introduction

Modern organizations build their competitive advantage on digital technologies and possibilities provided by data. To succeed in this kind of modern work life, students need various digital skills, including contemporary data analytics skills. Therefore, it is crucial to teach relevant methods, approaches, and tools for students to achieve contemporary skills for data analytics, but previous literature highlights several challenges that the educators face in teaching data analytics. Especially there is lack of innovative learning environments and different kinds of hands-on experiences that would transform the teaching of statistics and data analytics towards more applied and critical approaches, favoring the humanistic thinking of students.

Innovative learning environments offer also the possibility to apply learning analytics to enhance the learning process of the students (Sharif & Atif, 2024). Examining and measuring the data collected on learning platforms is learning analytics. Learning analytics includes various activities that help the teacher to support learning and optimize the learning environment. These activities include e.g. collecting, analyzing and reporting the learning data. Learning analytics data can be used to support the student and produce automatic feedback and adaptive tasks to guide the student in the right direction. Based on the data, the teacher is able to deepen the understanding of the learner's experience and develop teaching. The digital footprints of student activities left on learning platforms open the learning process and help monitor and analyze it in automatically. In this way the teacher can free up time for personal interaction with the student (Elias, 2011). However, previous literature has identified several different kinds of challenges faced in the adoption of learning analytics to teaching in higher education (Alzahrani et al., 2023; Gibson & Ifenthaler, 2020; Ferguson, 2012). To take the advantage of learning analytics in teaching technically demanding context such as data analytics, more empirically grounded information on practical trials is needed (see e.g. Wong, 2017).

In this paper, the challenges of teaching technical content, such as data analytics, in online course environment are studied. The special interest is to study how learning analytics can be used in teaching technical content for student for pedagogical development. The aim is to provide understanding both on the benefits and the pitfalls, and to develop potential solutions to overcome the challenges. The analysis of the usefulness of learning analytics is carried out from the teachers' perspective. Learning analytics experiences from two online data analytics courses are gathered in the form of teacher interviews.

Theoretical Background: What Are the Key Data Sources for Learning Analytics?

Learning analytics relies on various data sources to provide insights into student learning and optimize learning environments. The typical data sources for learning analytics include institutional datasets, digital traces, learning environment systems, Experience API (xAPI), and educational data specifications. Historically, learning analytics has relied on institutional datasets gathered from learning environment systems and registration data (Prinsloo et al., 2023). Digital traces that the various learning activities leave, such as log data, given answers, or source code, are nowadays valuable sources for learning analytics (Steinmaurer, 2021; Hoel & Chen, 2020).

Learning environment systems provide a wealth of data that is invaluable for learning analytics. One key type of data is completion rates, which show whether learners are

completing courses and how long it takes them to finish each task or module. This data helps in assessing the effectiveness of the eLearning course (Prinsloo et al., 2023). Learner performance and progress data includes information on how learners are performing in various tasks and their progress through the course. This data can highlight areas where learners are excelling or struggling, allowing for targeted interventions (Steinmaurer, 2021). Assessment scores from quizzes, tests, and other assessments provide measurable data on learner understanding and performance. This kind of learning data could help to identify both the weak and strong points of the student and to personalize further learning activities (Hoel & Chen, 2020).

Engagement metrics include data on how learners interact with the learning environment system, such as login frequency, time spent on different activities, and participation in discussion forums. These metrics provide insights into learner engagement and motivation (Wibawa et al., 2021). Digital traces, such as log data, answers given, and other digital footprints left by learners as they interact with the learning environment system, can be analyzed to understand learning behaviors and patterns (Steinmaurer, 2021). Interaction data includes information on interactions with other learning environment system features, such as discussion forums, peer reviews, and collaborative projects. This data provides insights into social learning and peer engagement (Hoel & Chen, 2020). Finally, search terms used by learners within the learning environment system can indicate areas of interest or confusion, helping to refine content and resources (Wibawa et al., 2021).

By leveraging these data points, educators and administrators can deepen their knowledge of the learning process, pinpoint areas that should be developed, and further, create more effective and personalized learning experiences. Survey responses from learners offer direct insights into their experiences and satisfaction with the course. This qualitative data is crucial for understanding the learner's perspective and improving course design (Prinsloo et al., 2023).

The use of a common standard like Experience API to gather activity data of the learning from multiple sources can facilitate data integration for learning analytics (Samuelsen et al., 2019, 2021). Furthermore, challenges such as data sharing, interoperability, and ethical implications need to be addressed to effectively utilize these data sources for learning analytics.

Research Methods and Data Sources

Learning analytics experiences from two online data analytics courses were gathered in the form of teacher interviews. Thus, the analysis of the usefulness of learning analytics is carried out from the teachers' perspective. These two courses were provided by South-Eastern Finland University of Applied Sciences, as part of a larger Knowledge-based management and data analytics open university curriculum provided as cooperation between three Finnish universities of applied sciences. This 25 credit unit curriculum included altogether six courses:
- Knowledge Management goals, 5 credit units (provided by Lapland University of Applied Sciences)
- Data-driven operations and development, 5 credit units (provided by Lapland University of Applied Sciences)
- Data analytics of management with information, 5 credit units (provided by South-Eastern Finland University of Applied Sciences)

- Tools and methods for data analysis, 5 credit units (provided by South-Eastern Finland University of Applied Sciences)
- Information management in organizations, 5 credit units (provided by Karelia University of Applied Sciences)
- Information Economy, 5 credit units (provided by Karelia University of Applied Sciences)

In this paper, the focus is on the two data analytics courses "Data analytics of management with information" and "Tools and methods for data analysis", provided by South-Eastern Finland University of Applied Sciences and the learning analytics used in these two courses. These two data analytics courses included a lot of technical content, such as SQL and R language, use of the Power BI program, and infographics. Thus, they were rather demanding for the students, but also offered many kinds of possibilities to include learning analytics solutions and to try automation of the course assignments. This makes these two courses interesting to study from the viewpoint of learning analytics.

The qualitative interviews focused on the use of learning analytics in these two online courses of data analytics, discussing available analytics, their benefits, challenges, and suggestions for improvement. Interview questions are presented in Table 1.

Table 1: Interview Questions for the Teachers

Interview questions for the teachers
What kind of learning analytics was available for these courses and what could have been available?
What should have been done differently in these courses in the use of learning analytics?
What does it take to utilize learning analytics in any course?
Based on your experience, what are the benefits of learning analytics for teachers? What benefits do you see learning analytics for students?
How to share learning analytics results with students a) during the course, b) after the course
What does the teacher do after receiving the results, and how could the results be used sensibly?
What data sources of the following a) SQL queries b) Panopto c) questionnaires d) Moodle have you used in these courses? What other possible data sources could be used?
On what basis did you choose these data sources?
Tell us about each data source you use - How have you used the information? - How do you assess the reliability of data? - How easily data has been available? - What were the challenges? - How useful it has been / what kinds of benefits you have gained as a teacher?
What would your dream learning analytics be like?

According to the interviews, the main data sources for learning analytics during these online courses were Moodle, Panopto, the SQL server and the student feedback surveys. Below we have a brief introduction to these data sources, but the key findings are presented in the next section.

Moodle is an open-source learning management system used in the above-mentioned online data analytics courses. In generally, it is commonly used learning management system, at least in Finland. On the Moodle platform of a course, a teacher can, for instance, share

documents and videos. In addition, it is possible to add assignments, exams and conversation areas for students there.

Panopto is a platform for creating and storing videos. Utilizing Panopto it is possible to get data on students' viewing activity concerning shared videos. This is important since videos are a key part of the online courses. Data on the SQL server is obtained only in the special case where students use SQL program language. More precisely, the data on SQL queries (made by students in Azure Data Studio) was collected from Azure SQL server. The student feedback surveys collecting basic information on students' opinions about the courses were conducted at the end of the courses.

Key Empirical Findings

Following Table 2 summarizes the main themes and insights from the interviews, highlighting the importance of structured planning and effective use of learning analytics in higher education.

Table 2: Key Findings From the Teacher Interviews

Data source	Useful information	Data reliability	Challenges	Benefits
Moodle	Basic information such as when students are on the platform, what they are doing there and for how long You can follow the progress of the ready-made process. How many people actually do tasks and how the number develops over time The exams already provide quite detailed information about the competence of individual students in the exam	Data reliable and available without doing anything yourself	Do-it-yourself analytics in Moodle is quite laborious Requires time from the teacher to dig into this information It only looks at activity that has taken place in Moodle, thus it doesn't tell how much time student has spent in total in doing the course	Familiar tool for many teachers, so there is no need to learn new things Includes basic analytics

Panopto	Information on the number of views of the videos and the time spent on them	The data itself is reliable, with some limitations to be taken into account	It only offers the results of all students' viewing in aggregate form if teacher does not use Panopto's log to open up the viewing of individual students' videos Videos can technically be viewed by all students, regardless of whether they are on the course or not -> affects the aggregated data	Panopto already has basic analytics
SQL server	Analytics examines the correctness of the syntax of the sentence written by the student	The SQL server collects the data correctly, there are no errors	Does not in itself reveal whether the answer is logically correct	Potential to build functional analytics around the data source, which would quite effectively monitor the training Students also has access to the correct answer to the SQL statement so that they could check more than just the correctness of the syntax
Student feedback surveys	Student perceptions on the course and its implementation etc If open questions included, may provide useful development ideas	The data is as reliable as it usually is in surveys Limitations related to the sample size - it is not worth drawing very far-reaching conclusions from the results	Too extensive questions may decrease the amount of answers The usefulness of a feedback survey largely depends on how successful the questions are In an open answer, the comment is always an individual person, so what is its weight for a teacher to develop the course	The feedback is available to the teacher almost without their own input especially if there are ready made survey questions provided by the organization

As seen from Table 2, learning analytics data was primarily collected from the Moodle platform, which tracked student activity within it. Video engagement was monitored through Panopto; however, the lack of mandatory logins led to potential inaccuracies in identifying viewers. Additionally, SQL server logs provided insights into student interactions with course materials.

For teachers, learning analytics assist in course development by identifying content that is either ineffective or overly challenging. For students, these analytics offer insights into their progress, helping them understand what is required to achieve their desired grades. The absence of clear goals for the use of analytics resulted in vague implementations. Concerns regarding data protection limited the ability to combine various data sources, preventing a comprehensive view of student performance. Furthermore, the reliance on Moodle's analytics, which only capture in-platform activities, may not accurately reflect overall student engagement.

It is recommended to establish clear objectives for what the analytics should achieve before course implementation. Incorporating real-time analytics would allow for timely interventions and adjustments. Additionally, enhancing the communication of analytics results to students by focusing on actionable insights rather than overwhelming data would be beneficial. The ideal learning analytics system would provide real-time feedback and actionable insights for both students and teachers, thereby enhancing the learning experience and the effectiveness of the course.

Conclusions and Discussion

In this paper, the interest was to study how learning analytics can be used in the courses of data analytics utilizing Power BI, the R and SQL program languages. We focused on the data sources. The aim was to provide understanding both on the benefits and challenges of the data sources in the sense of learning analytics. The analysis of the usefulness was carried out from the teachers' perspective. Furthermore, we focused on data and data sources, but it is important to utilize results in future, and measure how the changes made effect. In the best case, this would lead to a continuous process.

There is no doubt that learning analytics can bring several benefits that help the teachers to develop their teaching processes and support the students in their learning process. However, to really succeed in harnessing learning analytics for value creation in higher education, the higher education institutes should encourage teachers to adopt learning analytics through strategic approach. Thorough strategic approach the higher education institutions could create an environment where teachers feel empowered and motivated to utilize learning analytics, ultimately enhancing the educational experience for both educators and students. In next, some guidelines for this kind of strategic approach are presented.

Firstly, establishing of clear objectives for how learning analytics will enhance teaching and learning outcomes is crucial. Communicating the benefits of learning analytics by highlighting success stories and data showing improved student performance linked to analytics use can further support this vision. Secondly, by encouraging collaboration among teachers by creating communities of practice where they can share experiences and best practices is essential. Recognizing and rewarding teachers who effectively integrate analytics into their teaching can also foster a supportive culture. Additionally, peer mentoring can be implemented by pairing experienced users with those new to analytics for hands-on guidance.

Thirdly, regular training sessions, such as workshops and seminars, should be organized to familiarize teachers with analytics tools and their benefits. Fourthly, it is important to ensure that teachers have enough resources and also access to relevant software tools for effective analytics. Offering ongoing technical support to resolve any issues quickly can further facilitate the use of learning analytics by providing user-friendly analytics platforms that are easy to navigate is important.

In overall, teachers should be encouraged to consider analytics already during the course design phase to align teaching strategies with data insights. Also, addressing concerns about data security and privacy can help to build trust in using analytics. In future research it would be beneficial to study in more detail how teachers could apply the available learning analytics data and data sources in more effective way and how to better consider the students' point of view of learning analytics. These studies would need more empirical data gathered not only from teachers but also from students.

References

Alzahrani, A. S., Tsai, Y.-S., Iqbal, S., Marcos, P. M. M., Scheffel, M., Drachsler, H., Kloos, C. D., Aljohani, N., & Gasevic, D. (2023). Untangling connections between challenges in the adoption of learning analytics in higher education. *Education Information Technology,* 28, 4563–4595.

Elias, T. (2011). Learning analytics. *Learning*, *1*, 22.

Ferguson, R. (2012). Learning analytics: drivers, developments and challenges. *International Journal of Technology Enhanced Learning*, 4(5/6), 304–317.

Gibson, D. & Ifenthaler, D. (2020). Adoption of Learning Analytics. In: Ifenthaler, D., Gibson, D. (eds) Adoption of Data Analytics in Higher Education Learning and Teaching. Advances in Analytics for Learning and Teaching. Springer, Cham.

Hoel, T., & Chen, W. (2020). Data Sharing for Learning Analytics–designing conceptual artefacts and processes to foster interoperability. In *24th International Conference on Computers in Education* (Vol. 24).

Prinsloo, P., Khalil, M., & Slade, S. (2023). A Critical Consideration of the Ethical Implications in Learning Analytics as Data Ecology. In *European Conference on Technology Enhanced Learning* (pp. 371-382). Cham: Springer Nature Switzerland.

Samuelsen, J., Chen, W., & Wasson, B. (2019). Integrating multiple data sources for learning analytics—Review of literature. *Research and Practice in Technology Enhanced Learning*, 14, Article11.

Samuelsen, J., Chen, W., & Wasson, B. (2021). Enriching context descriptions for enhanced LA scalability: a case study. *Research and Practice in Technology Enhanced Learning*, *16*(1), 6.

Sharif, H. & Atif, A. (2024). The Evolving Classroom: How Learning Analytics Is Shaping the Future of Education and Feedback Mechanisms. *Education Sciences,* 14(2): 176.

Steinmaurer, A., Tilanthe, A. K., & Gütl, C. (2021). Designing and Developing a Learning Analytics Platform for the Coding Learning Game sCool. In *Interactive Mobile Communication, Technologies and Learning* (pp. 547-558). Cham: Springer International Publishing.

Wibawa, B., Siregar, J. S., Asrorie, D. A., & Syakdiyah, H. (2021, April). Learning analytic and educational data mining for learning science and technology. In *AIP conference proceedings.* (Vol. 2331, No. 1). AIP Publishing.

Wong, B. T. M. (2017). Learning analytics in higher education: an analysis of case studies. *Asian Association of Open Universities Journal*, 12(1), 21-40.