Data Visualization for Learning Analytics Indicators in Programming Teaching

Ranieri Alves dos Santos
Graduate Program in Knowledge Engineering and Management, UFSC, Florianópolis, Brazil

Dalner Barbi
Graduate Program in Knowledge Engineering and Management, UFSC, Florianópolis, Brazil

Vinicius Faria Culmant Ramos
Graduate Program in Knowledge Engineering and Management, UFSC, Florianópolis, Brazil

Fernando Alvaro Ostuni Gauthier
Graduate Program in Knowledge Engineering and Management, UFSC, Florianópolis, Brazil

Abstract

Learning Analytics (LA) has the potential to transform the way we learn, work and live our lives. To reach its potential, it must be clearly defined, incorporated into institutional teaching-learning strategies and processes and practices. The main goal of this study is to list indicators to be used in learning analytics in programming teaching and how to expose their views. For the development of the indicator model, this study based on a qualitative analysis, using data visualization and business intelligence tools, in projects focused on Learning Analytics. As a result, four main indicators were mapped: accesses to the system, resources accessed, activities carried out and, performance in activities.

2012 ACM Subject Classification Information systems → Data analytics

Keywords and phrases learning analytics, data visualization, learning indicators

Digital Object Identifier 10.4230/OASIcs.ICPEC.2023.10

Category Short Paper

Funding This work was carried out with the support of the Coordination for the Improvement of Higher Education Personnel – Brazil (CAPES) – Financing Code 001.

1 Introduction

In the context of the academic domain, Learning Analytics (LA) has grown rapidly, with a large volume of special editions of scientific journals in education, psychology, computing and social sciences focused on this topic. Due to its recent use, the organization of its fundamental concepts, theories, techniques, methods, approaches, applications and strategies has been structured for its development in several fields of action.

Since the emergence of digital teaching and learning platforms, which support educational processes, it has been possible to monitor the behavior of students and teachers in a broad and detailed way, offering a range of precision and prediction to the effectiveness of learning. And, in this context, new computational and interactional techniques are being created to absorb all the issues that involve it, from statistical results, purely, to issues such as ethics and privacy [17, 4, 11].

The field of Learning Analytics is attractive as a result of the possibility of using large volumes and a variety of data that enhance the vision of substantial improvements in teaching and learning practices, in an optimized and scalable way [9, 12].

As Clow [4] exposes, the exponential availability of data on learning activities in digital environments, with an approach centered on quantitative metrics, contributes to the understanding of educational processes that, in the context of the analysis, offers a broad view for teachers to use its resources, conceptions and teaching methods more effectively, providing better results.

Indeed, the field of Learning Analytics has been developing in a society permeated by the power of algorithms, mathematical approaches, advanced conditions of technological information processing equipment and data analysis methods, impacting education management and the process assessment as a foundation for the epistemological assumptions of pedagogical practices [8, 6].

2 Indicators for Learning Analytics

As the field of Learning Analytics is very recent, indicators of learning behavior in virtual learning environments with some degree of standardization have not yet been determined, to be used in the development of learning analyzes [12, 16].

Using a group concept mapping (GCM) approach, a tool available online, [16] create five dimensions of qualitative indicators for use in Learning Analytics assessments: a) objectives (awareness, reflection, motivation, behavioral change); b) learning support (perceived usefulness, recommendation, classification of activities, detection of students at risk); c) measures of learning and outcomes (comparability, effectiveness, efficiency, usefulness); d) data aspects (transparency, data standards, data ownership, privacy); and, e) organizational aspects (availability, implementation, training of educational stakeholders, organizational change).

In another study, Maraza-Quispe et al. [12] use a quantitative approach to propose indicators of learning behavior in virtual environments, to efficiently develop analytical learning processes, aiming at more effective predictions about learners' performance, decision-making and optimization of learning processes. They are: a) preparation for learning; b) progress along the course; c) learning resources; d) interaction in the forums; and, e) assessment of resources.

3 Data Visualization

Data visualization aims to provide ways to graphically represent collections of data. There are several ways to visualize and represent data, which go back to the first forms of information representation, such as maps, graphs and other graphics[18]. Data visualization technology acts as a cognitive aid to understand what you want to report, and it is a device that allows you to visualize nuances laid out only in the form of raw data in an understandable way. In this way, these techniques facilitate research, add meaning, capture information, collaborative discoveries and assist in the process of discovery and identification of elements.

Although clarity, objectivity and the ability to allow logical reasoning between data represent an aspect of data visualization, the reliability of the database and the reality it represents has a clear value relationship with visualization capabilities [13, 2, 3]. However, the data cannot speak for themselves, requiring the interlocutors to articulate variables for them in the discourse, from which the expected inferences of the reader can be generated. This work, sometimes attributed to data journalists, sometimes to researchers, condenses, aggregates, cuts or adjusts data so that it can be analyzed and disseminated. According to his ongoing project, with the advent of big data and the democratization of data visualization, it has become essential that executives and business analysts have the power to choose the information that is relevant to them [18].

The neutrality of the graphical representation agents as authors of the information and the neutrality of the elements that previously develop data acquisition patterns and variables are usually analyzed as much as the information provided by these agents. Therefore, just as the possibility of improving the comprehensibility of quantitative information through cognitive sources is widely discussed, it is essential to discuss the magnitude of biases that can be aggregated within the same source of information and the aspects that affect its reliability in projects of data visualization [2].

New ways of presenting data make it possible to instantly find insights that can now be displayed anywhere in the world. Thus, new data visualization tools and techniques address an old problem where knowledge is seen as synonymous with power, but having too much information can be counterproductive [18].

4 Methodological Procedures

For the development of the indicators model, this work was based on a qualitative analysis of approaches that use data visualization and business intelligence in projects focused on Learning Analytics. Therefore, a scope review was carried out using the acronym PCC (Population, Concept and Context).

The population was stipulated as being “indicators in Learning Analytics”, the concept as “data visualization techniques” and the context as “technologies in higher education”. As a research period, references published between 2012 and 2023 were sought, performing the search with the string “learning analytics” AND (“data visualization” OR “indicators”). The research sought to integrate several sources of publications about educational technology, for which the search used the LearnTechLib, Google Scholar and Web of Science databases and resulted in 1655 records.

Considering the only exclusion criteria is articles not related to the research question of this work, an exploratory qualitative research was carried out, focusing on the mapping of initiatives in Learning Analytics and data visualization. All articles that did not present relevant proposals on the use of visualization of learning analytics data and that were applicable to the teaching environment of programming in e-learning were excluded, in the end, 7 articles were selected.

5 Results and Discussions

From the works resulting from the review, Ali et al. [1] propose a tool for visualizing qualitative data in learning analytics, Essa and Ayad [5] propose a student success analysis system based on the visualization and prediction of learning data analytics, meanwhile, Khuzairi et al. [7] reviews the use of data visualization in learning analytics in the literature, Paiva et al. [14] proposes data mining in learning analytics, with data visualization, Phillips et al. [15] develops a classification algorithm for student engagement data in learning analytics, Song [19] analyzes data mining and visualization techniques in the context of learning analytics, Uskov et al. [20] works on the analysis of learning in higher education using data cleaning and visualization techniques.

The indicators to compose the dashboard proposed in this work were based on the works of Khuzairi et al. [7], Phillips et al. [15], Essa and Ayad [5] and Ali et al. [1]. These works served as a reference for defining the variables that will serve as a basis for viewing and classifying the educational data that will be displayed on the dashboard. Based on the works by Song [19], Paiva et al. [14] and Khuzairi et al. [7], data visualization strategies and navigation modes between data contexts (data drilling) were developed.

As a result of the research, four main indicators were mapped, raised in the literature that would be able to provide visualizations that would facilitate the process of learning analytics in programming teaching. Accesses to the System: indicator that measures the number of accesses to the system that the user made, regardless of the actions carried out in the educational system, only their accesses.

Resources Accessed: indicator referring to the amount of resources accessed within the educational system, compared to the amount of resources available.

Performed Activities: indicator that measures the amount of activities carried out in the system, comparing with the amount of activities available, without taking into account their mistakes and successes, only the performance.

Performance in Activities: indicator that measures errors and successes in activities available in the system.

The indicators of "Accesses to the System", "Resources Accessed" and "Performed Activities" are important to measure the performance of programming students because, according to Phillips et al[15], students with too much access to educational resources tend to be very effective in balancing your study commitments, contributing to your performance. In this way, on the contrary, when the student does not access the educational resources on the e-learning platform, his programming learning is compromised. The act of accessing or not accessing educational resources and the platform is not the main indicator of success or failure of the programming student. For Macarini et al[10], for example, counting any student interaction with the platform may be enough to predict early failure of programming students.

However, it is an indicator that contributes to the visualization of the learning analytics data about the student's commitment. "Performance in Activities" is the main indicator of student performance success to be visualized. This indicator, along with the three other indicators, make up the gauge visualization strategy in the proposed dashboard.

Based on the mapped indicators, a proposal for data visualization of these indicators was developed based on the "BI Dashboard" visualization style, with contextual visualizations, enabling data drilling by class and by students. Visualizations will be plotted in bar, pie and line graphs (Figure 1).

On the upper left side, the first data drill option appears, where the user can access views by class, in the example of Figure 1, the performance of the indicators in the class of "Object-Oriented Programming – 23948203-1" is displayed. The teacher will be able to browse the information of their classes and the educational manager will be able to access the data of all classes.

Just below, on the left side, the second data drill option is displayed, which allows the teacher to choose which student he would like to visualize his performance. The teacher or manager can group one or more students to obtain their average performance.

In the visualization graphs, it will be possible to observe the student's performance in the activities carried out, where from a bar graph, their assertiveness can be measured against the expected context. In the next pie chart, the teacher will be able to easily visualize the percentage of available resources that the user has accessed, the same happens with the next pie chart, which visualizes the percentage of activities carried out by the student. Just below, the line graph displays the number of accesses to the system in the historical series and a gauge with the student's classification.

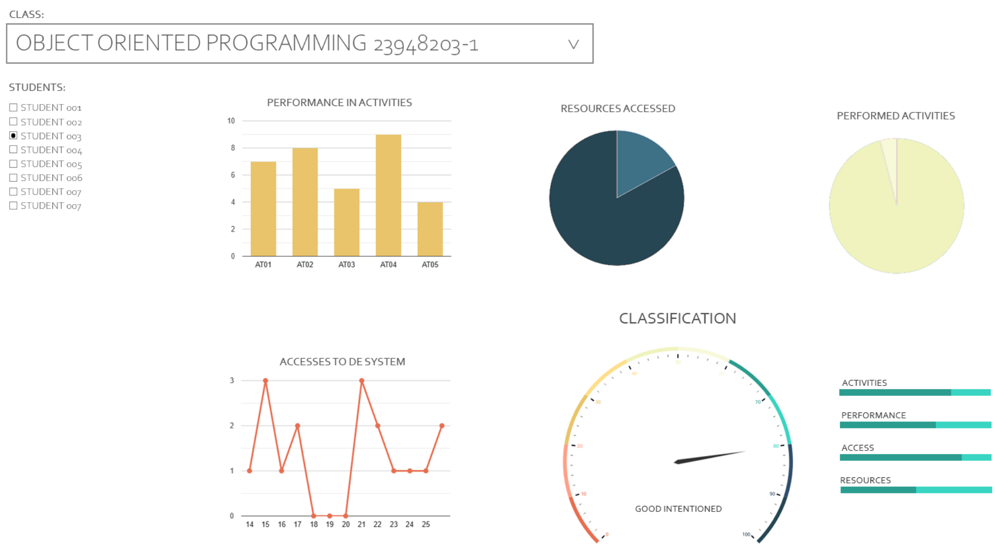

Figure 1 Overview of the Proposed Model Dashboard.

The student rating scale is displayed in a gauge at the lower right corner (Figure 2). The gauge is based on the algorithm by Phillips et al [15], which uses the four indicators mapped in this work to specify the student's level of involvement within the course. This algorithm, based on the indicators, classifies the student in ascending order in the following terms: non-user, random, single-use, accidental, free-time, inserted, constant user, well-intentioned, engaged and conscious, the latter being the student with the highest ranking against the algorithm.

Figure 2 Overview of the Proposed Model Dashboard.

6 Conclusion

With the development of the proposed Learning Analytics data visualization model applied to programming teaching, it was possible to observe that the literature presents several approaches that facilitate the development of new approaches for the analysis of educational data from the actions already carried out and mapped fundamentally in online educational and learning management systems.

In this way, the present work presented a form of data visualization based on existing approaches in the literature capable of facilitating the analysis of the performance of classes and of specific students from contextual visualizations with data drilling navigation.

Initiatives aimed at simplifying quantitative symbologies from raw data in infographics, which are capable of composing strategies that facilitate the understanding of information for the appropriate target audience, can be of great contribution to the most diverse sectors, especially in education. Teachers and educational managers having access to simplified graphical views, which allow a quick understanding of the current status of certain students or classes, can dedicate more time to other activities, since they will no longer need to look for reports of grades, accesses and activities on several screens, since that a dashboard like the one proposed here is able to quickly inform those involved about the current performance of the selected sample.

As a way of continuing the work developed based on this, it is suggested that further research be carried out on the proposed model, implementing its dashboard in programming teaching institutions that have systems capable of retrieving information on indicators from educational databases.

References

1 Liaqat Ali, Marek Hatala, Dragan Gasevic, and Jelena Jovanovic. A qualitative evaluation of evolution of a learning analytics tool. *Comput. Educ.*, 58(1):470–489, 2012. doi:10.1016/j.compedu.2011.08.030.

2 Jaqueline Vasconcelos Braga, Tiago Barros Pontes, Virginia Tiradentes Souto, et al. Statistical manipulations and visual anomalies: data visualization design and statistical bias recognition/manipulacoes estatisticas e anomalias visuais: design de visualizacao de dados e reconhecimento de vieses estatisticos. *Brazilian Journal of Information Design*, 17(2):145–163, 2020.

3 Alberto Cairo. *The truthful art: Data, charts, and maps for communication.* New Riders, 2016.

4 Doug Clow. An overview of learning analytics. *Teaching in Higher Education*, 18(6):683–695, 2013.

5 Abdullah Essa and Hany Ayad. Student success system: risk analytics and data visualization using ensembles of predictive models. In *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge*, pages 158–161. ACM, April 2012.

6 HU Hoppe. Computational methods for the analysis of learning and knowledge building communities. In *Handbook of Learning Analytics*, chapter 2. Society for Learning Analytics Research, 2017.

7 Nur Muizzatul Shafiqah Khuzairi, Zainal Choy Cob, and Tuan Hilaluddin. Towards understanding the synergetic relationship of data visualization with learning analytics: A review. In *AIP Conference Proceedings*, volume 2644(1), page 030030. AIP Publishing LLC, November 2022.

8 Simon Knight and Simon Buckingham Shum. Theory and learning analytics. In *Handbook of Learning Analytics*, pages 17–22. Society for Learning Analytics Research, 2017.

9 Vitomir Kovanović, Srećko Joksimović, Dragan Gašević, Marek Hatala, and George Siemens. Content analytics: The definition, scope, and an overview of published research. In *Handbook of Learning Analytics and Educational Data Mining*, pages 77–92. Society for Learning Analytics Research, 2017.

10 Lucas A B Macarini, Cristian Cechinel, Marilde F Batista Machado, Vanessa Faria Culmant Ramos, and Rodrigo Munoz. Predicting students success in blended learning—evaluating different interactions inside learning management systems. *Applied Sciences*, 9(24):5523, 2019.

11 Katerina Mangaroska and Michail Giannakos. Learning analytics for learning design: A systematic literature review of analytics-driven design to enhance learning. *IEEE Transactions on Learning Technologies*, 12(4):516–534, 2018.

12 Basilio Maraza-Quispe, Omar Marcelo Alejandro-Oviedo, Wilson Choquehuanca-Quispe, Nilton Caytuiro-Silva, and Juan Herrera-Quispe. Towards a standardization of learning behavior indicators in virtual environments. *International Journal of Advanced Computer Science and Applications*, 11(11), 2020.

13 Isabel Meirelles. *Design for information: an introduction to the histories, theories, and best practices behind effective information visualizations.* Rockport publishers, 2013.

14 Rodrigo Paiva, II Bittencourt, Wagner Lemos, Andrade Vinicius, and Diego Dermeval. Visualizing learning analytics and educational data mining outputs. In *International Conference on Artificial Intelligence in Education*, pages 251–256. Springer, June 2018.

15 Rob Phillips, Dorit Maor, Greg Preston, and Wendi Cumming-Potvin. Exploring learning analytics as indicators of study behavior. In *EdMedia+ Innovate Learning*, pages 2861–2867. Association for the Advancement of Computing in Education (AACE), 2012.

16 Maren Scheffel, Hendrik Drachsler, Slavi Stoyanov, and Marcus Specht. Quality indicators for learning analytics. *Journal of Educational Technology and Society*, 17(4):117–132, 2014.

17 George Siemens. Learning analytics: The emergence of a discipline. *American Behavioral Scientist*, 57(10):1380–1400, 2013.

18 Felipe Cezar Cardoso Da Silva. Data visualization: past, present and future. *LIINC em Revista*, 15(2):205–223, 2019.

19 Dongmin Song. Learning analytics as an educational research approach. *International Journal of Multiple Research Approaches*, 10(1):102–111, 2018.

20 Vladimir L Uskov, Jon P Bakken, Karthik S Ganapathi, Kevin Gayke, Blake Galloway, and Johra Fatima. Data cleaning and data visualization systems for learning analytics. In *Smart Education and e-Learning 2020*, pages 183–197. Springer, 2020.