Data literacy and learning analytics: exploiting the data of Students' assessment to support the learning of Mathematics in upper secondary school

Research project – Supervisor Prof. Stefania Mignani

Brief introduction

Data science impacts many aspects of our life. It has been transforming industries, healthcare, as well as other sciences. Education is not an exception. In all educational environments, one of the main objectives is to ensure that learning processes allow to understand students and their learning paths. An explosion of available data has revolutionized how much education research is done. Educational Data Mining (EDM) provide a fundamental value to educational institutions and to all entities that support different processes in learning activities.

Learning analytics (Analysis of Learning) is one of the new young disciplines in the area. LA can be defined as the measurement, collection, analysis and reporting of data about students and their contexts, for understanding and learning purposes and optimizing the environments in which occurs. Learning Analytics sits at the convergence of Learning (e.g., educational research, learning and assessment sciences, educational technology), Analytics (e.g., statistics, visualization, computer/data sciences, artificial intelligence), and Human-Centred Design (e.g., usability, participatory design, sociotechnical systems thinking). Motivated by the challenge to provide educators and learners with insights that might improve learning processes and teaching practice, the research project focuses on the application of statistical tools for classifying students in groups to promote peer activities. Biggs and Tang (2011) emphasize the importance of 'constructive alignment' between learning outcomes, learning activities design and assessment tasks in which these components support each other and put students in a supportive learning system.

Background and statement of the problem

The increasing interest in data use in education is twofold. On the one hand, there is the accountability context in which school leaders and teachers are held accountable for the quality of the education they provide (Lai & Schildkamp, 2013). On the other hand, there is a growing recognition that data should be for continuous improvement (Kingston & Nash, 2011; Lai & Schildkamp, 2013; Mandinach, 2012). In this context, data use is seen as a way to inform teachers about students' needs and adapt and adjust instruction based on such information.

Teachers collect information about their students all the time to help them make informed decisions. Technological developments enable educators to collect, analyze, interpret, and distribute data in increasingly efficient and systematic ways (Mandinach, 2012).

Over the last decade, learning analytics have received considerable attention in non-formal and formal learning contexts due to their potential for monitoring, supporting, assessing, and managing student learning. Wise and Jung (2019) conceptualized the process of teachers' use of LA highlighting how LA are used to revise the learning design process, to undertake scaffolding actions, and could be used as a reflective tool about their instructional intervention.
Wise (2014) emphasized the importance of LA in the educational context, introducing four principles of instructional design with the function of supporting, engaging, and empowering students toward their own learning process: Integration, Agency, Reference frame, and Dialogue.
Knight and al. (2017) represent through a triangle three dimensions Epistemology-Assessment-Pedagogy (EPA) to explain inter-relationships among them. Following the model, the authors argue that LA are able to support educational practices and reshape their approach.

The evidence from research and practice shows that there are far more productive and potent ways of using analytics for supporting teaching and learning. Some of the most popular goal of learning analytics include:

1. Supporting student development of lifelong learning skills and strategies
2. Provision of personalised and timely feedback to students regarding their learning
3. Supporting development of important skills such as collaboration, critical thinking,
4. Develop student awareness by supporting self-reflection
5. Support quality learning and teaching by providing empirical evidence on the success of pedagogical innovations

Learning analytics have made possible data-driven decisions for improving student learning, utilizing the data already collected by educational tools and platforms.
With the broader adoption of educational technologies in primary and secondary education and the emergence of new classroom-focused technologies (Voogt et al., 2018), there has been a growing awareness of the potentials of learning analytics for supporting students and diagnosing their learning progress in pre-university contexts.

Research questions or hypothesis, aim, objectives and deliveries

The Project is divided into two connected parts. The first part is addressed directly to the teachers and the second one is dedicated to the students.

With regard to teachers, the project aims to improve the skills of Data Literacy, understood as the ability to use quantitatively all the information about students' performance.

Processing the available data with statistical tools both promotes awareness of the learning environment and supports educational initiatives both to support students at risk and to value students deserving.

On the student side, actions aimed at learning and behavioural data can encourage greater involvement in the performance evaluation phase and encourage participation in specific activities even among peers.

The objectives of the project are:

1) Increasing the data literacy of teachers
2) Identify appropriate statistical techniques to support the classical assessment with an approach that follows the learning path of the student over time
3) Evaluate over time the impact on student performance
4) Implement a procedure that allows the teacher to promote peer-to-peer activities based on the data analysis available.

The project envisages carrying out a pilot case study with a focus on mathematics and as a reference for secondary students.

The research question is to define on the basis of the data available from electronic registers a dynamic clustering procedure to follow over time students' performance changes. The cluster analysis must identify groups of similar students who will have to carry out learning activities, both in formal and informal ways, where knowledge and skills are acquired through active assistance and support among each other (Nawaz and Rehman, 2017).

The pilot study is part of a collaboration under the PLS (Piano Lauree Scientifiche) with several secondary schools.

The teachers who participated in the Data Literacy initiative made themselves available to provide (under privacy protocol) data on the learning results of their students. Their involvement will allow to test the different solutions of data collection and implementation of an analysis of the groups with the aim of realizing for the scholastic year 22-23 peer learning activities to follow the changes in performances (Valarmathy, and Krishnaveni, 2019).

From a methodological point of view, the focus will be to deepen statistical tools for studying time-varying group (cluster) structures in panel data. Dynamic clustering continuously analyses this data to identify and predict changes in purchasing patterns. Moreover, there is a clear awareness among researchers and practitioners of the necessity to investigate novel approaches that are more

appropriate for many new clustering problems arising in diverse situations and requiring special treatment. The project will address the problem in the case of small data set (Den Teuling et. Al., 2021) On the basis of the results of this pilot study, it is intended to enlarge the group of participating schools.

Participants in the study and the role they play

The successful applicant will be part of the STAT department research group working on methods for educational evaluation (Stefania Mignani, Mariagiulia Matteucci and Matteo Farnè). The topic of Data Literacy will also be addressed involving both members of the Department (Angela Montanari; Silvia Bianconcini) and teachers from upper secondary schools.

Thanks to the agreement with the STAT department, some members of the Invalsi working on will collaborate to the research project by providing ad hoc data and giving support in the interpretation of results.

Moreover, international collaborations are active with Prof. Dr. Bernard P. Veldkamp (Department of Research Methodology, Measurement and Data Analysis, University of Twente, The Netherlands).

References

- Biggs, J., & Tang, C. (2011). *Teaching for Quality Learning at University*. Maidenhead, UK: Open University Press.
- Chang, W-C.; Wang, T-H.; Li, M-F. (2010). Learning Ability Clustering in Collaborative Learning,. *Journal of software*, vol. 5, no. 12,
- Kingston, N., & Nash, B. (2011). Formative Assessment: A Meta-Analysis and a Call for Research. *Educational Measurement: Issues and Practice,* vol. 30, pp. 28-37.
- Knight, S., & Buckingham Schum, S. (2017). Theory and learning analytics. In C. Lang, G. Siemens, A. F. Wise, D. Gašević (Eds.), The Handbook of Learning Analytics. (pp. 17-22). Alberta, CA: Society for Learning Analytics Research.
- Mandinach, Ellen B (2012) A Perfect Time for Data Use: Using Data-Driven Decision Making to Inform Practice, *Educational Psychologist*, vol. 47 n°2, pp. 71-85 2012
- Matteucci M., Mignani S. (2021), Investigating gender differences in mathematics by performance levels in the Italian school system, *Studies in Educational Evaluation*, Vol. 70 , pp. 1 - 12

- Nawaz, A.; Rehman Z. U. (2017). Strategy of Peer Tutoring and Students Success in Mathematics: An Analysis. *Journal of Research and Reflections in Education* June 2017, Vol. 11 , No 1, pp 15-30.
- Schildkamp, K., & Lai, M. K. (2013). Data-based decision making : conclusions and a data use framework. In K. Schildkamp, M. K. Lai, & L. Earl (Eds.), Data-based decision making in education: Challenges and opportunities . *Studies in educational leadership*; Vol. 17, n°. 17 pp. 177-191.
- Den Teuling, N. G. P.; Pauws, S. C.; van den Heuvel, E. R. (2021). A comparison of methods for clustering longitudinal data with slowly changing trends. *Communications in Statistics - Simulation and Computation*,
- Voogt, Joke & Knezek, Gerald & Christensen, Rhonda & Lai, Kwok-Wing. (2018). Second *Handbook of Information Technology in Primary and Secondary Education*. 10.1007/978-3-319-71054-9.
- Valarmathy, N.; Krishnaveni, S. (2019). Performance Evaluation and Comparison of Clustering Algorithms used in Educational Data Mining. *International Journal of Recent Technology and Engineering* (IJRTE) ISSN: 2277-3878, Volume-7, Issue-6S5.
- Vellido, A., Castro, F., Nebot, A. (2011). *Clustering Educational Data*. In Handbook of Educational Data Mining, Chapman & Hall, pp. 75-92.

- Wise, A. F. (2014). Designing pedagogical interventions to support student use of learning analytics. In A. Pardo & S. D. Teasley (Eds.), Proceedings of the International Conference on Learning Analytics and Knowledge. New York, NY: ACM Press.
- Wise, A. F., & Jung, Y. (2019). Teaching with analytics: towards a situated model of instructional decision-making. Journal of Learning Analytics, 6(2), 53–69.

Activity plan

The project is structured in the following steps:

1) Review of the literature about learning analytics and clustering
2) Collect and data Arrangement data from the students of schools involving in the project
3) Promote specific activities with teachers to increase data literacy using students' evaluation results
4) Implement the clustering procedure at start time
5) Implement dynamic clustering after the peer learning activities