

# Applying Process Mining Techniques to Learning Management Systems for Educational Process Model Discovery and Analysis

Darko Etinger<sup>1</sup>(✉), Tihomir Orehovački<sup>1</sup>, and Snježana Babić<sup>2</sup>

<sup>1</sup> Department of Information and Communication Technologies, Juraj Dobrila University of Pula, Zagrebačka 30, 52100 Pula, Croatia

{darko.etinge, tihomir.orehovacki}@unipu.hr

<sup>2</sup> Polytechnic of Rijeka, Trpimirova 2/V, 51000 Rijeka, Croatia  
snjezana.babic@veleri.hr

**Abstract.** Based on the event logs gathered from the learning management system in use at the Juraj Dobrila University of Pula, this study attempts to discover frequent behavioural patterns in event logs. The process mining tool Fluxicon Disco was used to create, clean and prepare the event log. The discovered process model represents the starting point for the assessment of the usage behaviour of the Moodle LMS. A procedure was developed in R (R-project) to analyse the process model by using various R based packages for business process analysis (bupaR, edeaR, processmapR). By applying process discovery, a subset technique of process mining, a process model is obtained, describing the usage behaviour based on real facts and evidence. This obtained educational process model can be used to tailor the Moodle LMS to the specific needs of the individual courses. Customized workflows for specific courses can be modelled based on the results of the analysis, improving the usability of the learning management system. The significance and importance of this study is achieved by applying process discovery as a rather new approach to improve the Moodle LMS usability with evidence and actual workflows, rather than to assess the learning management system usage by other techniques, e.g. based on users' perceptions.

**Keywords:** Educational process model · Process mining · Process discovery  
Learning Management Systems · Usability

## 1 Introduction

Educational process mining is an emerging field in the educational data mining discipline, concerned with discovering, analysing, and improving educational processes as a whole, based on information hidden in educational datasets and event logs [1]. Learning management systems provide facilities for managing the learning experience, communicating the intended learning experience and facilitating teachers' and students' involvement in that experience. In these systems, students can access courses' contents in different formats (text, image, sound, video), as well as interact with teachers and/or

colleagues, via message boards, forums, chats, video-conference or other types of communication tools. Learning management systems are widely implemented within Croatian universities with the scope of enhancing higher education. Moodle LMS, as a preferred learning management system at Juraj Dobrila University of Pula, is offered to students and teachers as a complementary tool to improve the teaching process and learning outcomes.

Process mining offers comprehensive sets of tools to provide fact-based insights and to support process improvements [2]. It combines Big Data, Data mining techniques and Process Modelling and Analysis. Process mining bridges the gap between traditional model-based process analysis (e.g., simulation and other business process management techniques) and data-centric analysis techniques such as machine learning and data mining [3]. Emerging from the field of Business Process Management (BPM), process mining focuses on extracting process-related knowledge from event logs recorded by an information system, in our case, the learning management system. By applying process mining to extract knowledge about the underlying processes in learning management systems, it is possible to analyse the usage behaviour according to the records in the event log. Data about each event contain a Case ID, a time stamp, an activity ID and description, and information about various resources. The main drivers for this technology are the omnipresence of event data that cannot be successfully analysed by traditional data mining tools and the underperforming of contemporary Business Process Management (BPM) and Business Intelligence (BI) software applications.

The remainder of this paper is organized as follows: theoretical foundation of our work is briefly described in the Sect. 2. It summarizes process mining techniques in the context of educational process mining. In Sect. 3, we explained the research methodology and presented the process mining and business process analysis results. Finally, contributions and implications of study results are outlined in the fourth section.

## 2 Theoretical Framework

Moodle LMS is a software package for producing Internet-based courses and web sites. It is a global development project designed to support a social constructionist framework of education, provided freely as Open Source software (under the GNU General Public License). It is the information system chosen to support the learning process at Juraj Dobrila University of Pula. As such, it is installed on a dedicated server and is maintained and administered by the IT support service.

Educational process mining (EPM) is an emerging field in the educational data mining (EDM) discipline, concerned with developing methods to better understand students' learning habits and the factors influencing their performance [4]. It aims at discovering, analysing, and providing a visual representation of complete educational processes [5]. Van der Aalst [6] noted that process mining may be used to discover the way that people really work, to find out where there are deviations, and for all kinds of process improvement. The results of EPM can be used to get a better understanding of the underlying educational processes, to generate recommendations and advice to students, to provide feedback to either students, teachers or/and researchers, to detect

learning difficulties early, to help students with specific learning disabilities, and to improve management of learning objects [4].

Van der Aalst [2] defined three types of process mining: discovery, conformance, and enhancement. Process discovery is one of the most challenging process mining tasks. Based on an event log, a process model is constructed thus capturing the behaviour seen in the log. It is a combination of two perspectives: discovery task and control-flow.

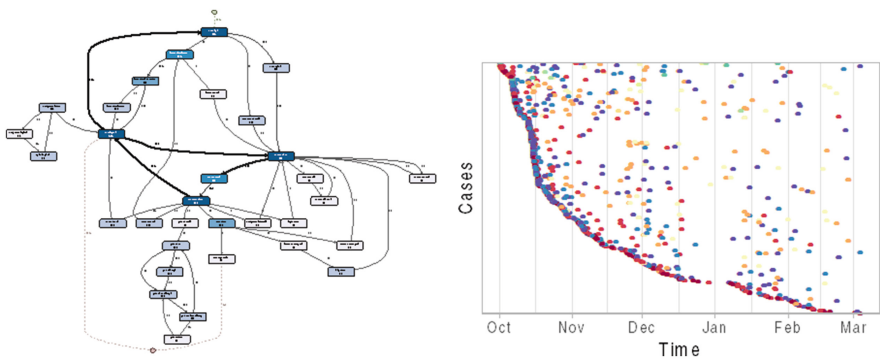
### 3 Research Methodology and Results

A large number of event data is contained in the Moodle LMS database, the learning management system in use at the Juraj Dobrila University of Pula. The event logs extracted from the Moodle LMS database cover a period of one semester, assuring that the time frame is relevant enough to bring insight into the learning process. The aim of applying process discovery, a subset technique of process mining, is to obtain a model that describes reality better than the pre-defined procedures and requirements.

Cairns et al. [5] noted that an event log is a hierarchically structured file with data about historical process executions, which is then used by process mining techniques. This file has to be constructed by structuring raw process data that can be found in files or databases (e.g., the Moodle LMS), into events and traces. An event is the most atomic part of a specific process execution, containing a name, a specific timestamp associated with the event, the originator of the event and other attributes. A trace is a collection of events that belong to the same process execution. A typical learning management system logs most of the user activities like courses attended, modules read, practice exams attempted, exam scores, students and teachers' interaction via chat logs or discussion boards, etc.

The analysis started by collecting all the relevant txt/csv files out of Moodle LMS. This raw data was then reorganized into an event log. Event logs were extracted from the Moodle LMS database, for the period ranging from October 1<sup>st</sup>, 2014 to March 4<sup>th</sup>, 2015, covering a complete winter semester. To transform the original data into a suitable shape to be used by process mining algorithms, the data collected from databases was reshaped into a consolidated log (stored as a CSV file and converted into an XES format file). Additionally, information about courses and specific resources accessed are collected. The event log (organized as a XES – eXtensible Event Stream) is then imported into Fluxicon Disco application to show in detail how the processes have been performed. To extract relevant information about the process model, several filters were applied. The process discovery and analysis consisted of two steps. First, we filtered cases containing the activities starting with “user login” and ending with “user logout”. A total of 434 cases were identified with 60 distinct atomic activities, comprising 159751 events with its own time stamp. In the second step, the fuzzy process model discovery algorithm was applied on the fragment of our event log containing users' courses over one semester, to get a big picture about the nature of the underlying process. A desired process map containing 19% of all cases was obtained (shown in Fig. 1). Janssensvillen et al. [7] developed ‘bupaR’ which consists of different R-packages, each with their own purpose: bupaR (basic functionality for handling event data), edeaR (for exploratory

and descriptive analysis of event data), processmapR (for creating process visualizations) and several others not used in this study. The dotted chart shows the spread of events over time by plotting a dot for each event in the log, allowing us to gain insight about the occurrences of the events over time in the complete set of data [8]. The chart has three (orthogonal) dimensions: the time of the event, the case ID and the occurrence of the event. The dotted chart for the discovered model was visualized with the R package ‘processmapR’ (as shown in Fig. 1).



**Fig. 1.** Educational process model and dotted chart discovered and visualized after filtering 19% of the cases containing the “user login – user logout” pattern

Cairns et al. [5] stated that event logs in the education domain, particularly those coming from e-learning environments, may contain massive amounts of fine granular events and process related data. The real-life process for Moodle LMS is very unstructured and traditional process mining approaches have problems dealing with unstructured processes. The discovered models are often “spaghetti-like” showing all details without distinguishing what is important and what is not [2]. Thus, to obtain a usable model, the goal was to discover a model that fits well with the general students’ behaviour, meaning it must not be too large nor too complex for the analysis.

Tax et al. [9] proposed the Local Process Models (LPMs) allowing the mining of patterns positioned in-between simple patterns and end-to-end models, focusing on a subset of the process activities and describing frequent patterns of behaviour. Thus, we analysed the model with ‘edeR’ to obtain a list of activities that are most frequent: user login (1), course view (2), user logout (3), resource view (10), course enrol (18), user view (19), forum - view forum (24), forum - view discussion (27), as shown in Fig. 2. Based on these activities, a new process model was filtered out in Fluxicon Disco, showing the most frequent usage behaviour patterns (as shown in Fig. 3.).

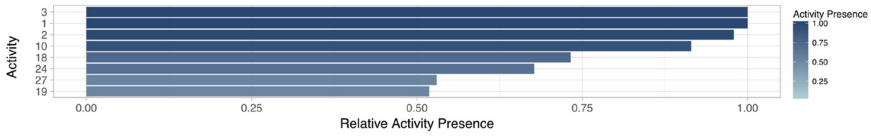


Fig. 2. Activity frequency chart (created with edeaR)

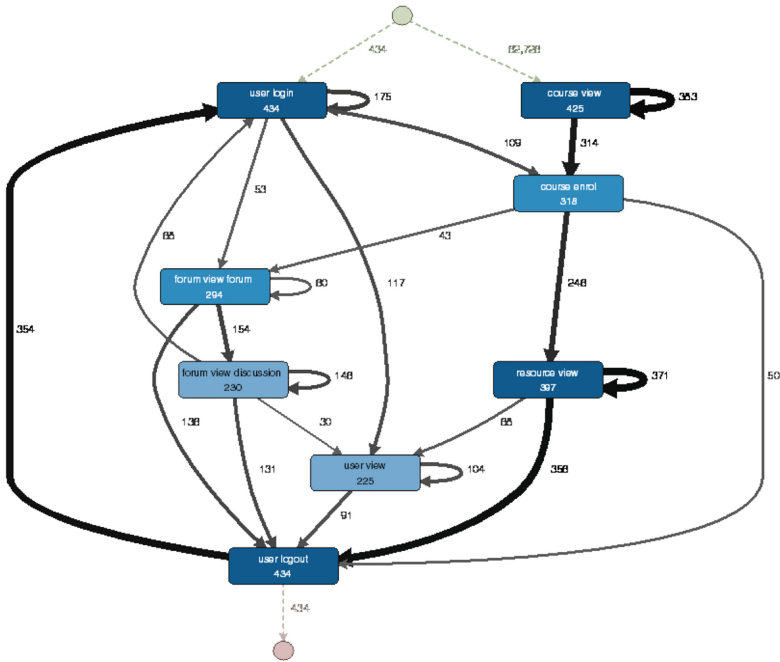


Fig. 3. Educational process model for Moodle LMS (created with Fluxicon Disco) with corresponding activity frequency chart (created with edeaR)

The final model describes the most common usage behaviour patterns by users of the Moodle LMS. It is evident that users mostly use forums for collaboration and access learning materials provided by the course teachers. As this model provides insight about usage behaviour based on real facts and evidence, it enables a better understanding of the underlying educational processes. This educational process model can be used to tailor the Moodle LMS to the specific needs of the individual courses, to generate recommendations and advice to students, to provide feedback to either students, teachers or/and researchers, to detect learning difficulties early, to help students with specific learning disabilities and finally, to improve management of learning objects.

## 4 Concluding Remarks

The idea of educational process mining is to detect, monitor and improve real-life processes by extracting process-related knowledge from learning management system event logs. As an emerging field in the educational data mining discipline, educational process mining allows to get a better understanding of the underlying educational processes. This study attempted to discover frequent behavioural patterns in event logs. By applying process discovery, a subset technique of process mining, the process model that describes the usage behaviour based on real facts and evidence was obtained. The adoption of filtering and abstraction techniques to reduce the complexity of the discovered process model enabled the extraction of an easy-to-use and valuable process model. By establishing a direct connection between the process models and event data, an “evidence-based” process improvement is enabled. Insight into the user experience with learning management systems enable universities to understand and define policies and actions to improve the usability of learning management systems and promote usage among users. Potential benefits may also include better students’ and teachers’ satisfaction and increased effectiveness of the learning process.

## References

1. Cairns, A.H., Gueni, B., Assu, J., Joubert, C., Khelifa, N.: Analyzing and improving educational process models using process mining techniques. In: *IMMM 2015 Fifth International Conference on Advances in Information Mining Management*, pp. 17–22 (2015)
2. Van der Aalst, W.M.P.: Trends in business process analysis: from verification to process mining. In: *Proceedings of 9th International Conference on Enterprise Information Systems (ICEIS 2007)*, pp. 12–22 (2006)
3. Mans, R.S., van der Aalst, W.M.P., Vanwersch, R.J.B.: Process Mining. In: *Process Mining in Healthcare: Evaluating and Exploiting Operational Healthcare Processes*, pp 17–26. Springer International Publishing, Heidelberg (2015)
4. Ariouat, H., Cairns, A.H., Barkaoui, K., Akoka, J., Khelifa, N.: A two-step clustering approach for improving educational process model discovery. In: *2016 IEEE 25th International Conference on Enabling Technologies Infrastructure for Collaborative Enterprises*, pp. 38–43 (2016)
5. Cairns, A.H., Gueni, B., Fhima, M., Cairns, A., David, S., Khelfa, N.: Process mining in the education domain. *Int. J. Adv. Intell. Syst.* **8**, 219–232 (2015)
6. Van der Aalst, W.M.P.: Business process management: a comprehensive survey. *ISRN Softw. Eng.* **2013**, 1–37 (2013)
7. Janssenswillen, G., Depaire, B.: BupaR: business process analysis in R. In: *CEUR Workshop Proceedings*, vol. 1920, pp. 2–6 (2017)
8. Pechenizkiy, M., Trcka, N., Vasilyeva, E., van der Aalst, W.M.P., De Bra, P.: Process mining online assessment data. In: *Proceedings of 2nd International Conference on Educational Data Mining*, pp. 279–288 (2009)
9. Tax, N., Sidorova, N., Haakma, R., van der Aalst, W.M.P.: Mining local process models. *J. Innov. Digit. Ecosyst.* **3**, 183–196 (2016)