

Educational Process Mining-Different Perspectives

K.Lalitha Devi M.Sc., M.Phil.,¹ M.Suryakala M.Sc., M.Phil., P.hD²

¹(Computer Science, M.R.Govt.Arts College, India.)

² (Computer Science, Dr.Ambethkar Govt.Arts College, India.)

Abstract: Process mining methods have in recent years enabled the development of more sophisticated Process models which represent and detect a broader range of student behaviors than was previously possible. This paper summarizes key Process mining perspectives that have supported student modeling efforts, discussing also the specific constructs that have been modeled with the use of educational process mining and key upcoming directions that are needed for educational process mining research to reach its full potential. Process mining aims to discover, monitor and improve real processes by extracting knowledge from event logs readily available in today's information systems. This paper is designed to give a view on the capabilities of process mining techniques in the context of higher education system which involves and deals with administrative and academic tasks like enrolment of students in a particular case, alienation of traditional classroom teaching model, detection of unfair means used in online examination, detection of abnormal values in the result sheet of students, prediction about students performance, identify the drop outs, and students who need special attention and allow the teacher to provide appropriate advising /counseling and so on.

Keywords: event logs, educational process mining, perspectives,

I. Introduction

Current education System is learner Centric. A teacher facilitates learning to the students, and for learning, the learners not only dependent on the lessons in the classroom teaching, but learns from internet, elearning softwares, and cds and from different digital content. ICT has made education system as learner centric. In the education system the learners interact with the intelligent information systems designed to educate them are capable of recording the interactions of the learners and they store enormous amount of records called logs. The learners are assessed not only with a skill or concept but a broad range of affective, meta cognitive, motivational and behavioral constructs. So the underlying implicit information that the logs contain and its potential use has created the necessity to develop tools, which are able to mine the information and make it explicit with the help of formal representation. Process mining is one such tool and which is sub domain of data mining. This technique use event logs to discover analyze and improve processes and process models. Processes mining technique take a set of real execution (the so-called event log) as the starting point and these techniques attempt to extract non-trivial and useful process information from various perspectives like control, data flow, organizational structure and performance perspective.

In this paper, we give a brief overview of Process mining, framework and its perspectives in higher educational domain to improve learner behavioral model. In educational environment similar to business information system many tasks are performed both from administration and academic activities. The information systems used in this environment record event logs, from which the knowledge discovery can be made to improve the process model.

II. Process Mining Framework

Process mining has emerged from the field of Business Process Management (BPM). It focuses on extracting process-related knowledge from event logs¹ recorded by an information system. It aims particularly at discovering or analyzing the complete (business, or in our case educational) process and is supported by powerful tools that allow getting a clear visual representation of the whole process. The three major types of process mining applications are (Figure 1):

- 1) Conformance checking - reflecting on the observed reality, i.e. checking whether the modeled behavior matches the observed behavior;
- 2) Process model discovery - constructing complete and compact process models able to reproduce the observed behavior, and
- 3) Process model extension - projection of information extracted from the logs onto the model, to make the tacit knowledge explicit and facilitate better understanding of the process model.

Process mining is supported by the powerful open-source framework ProM. This framework includes a vast number of different techniques for process discovery, conformance analysis and model extension, as well as many other tools like convertors, visualizers, etc. The ProM tool is frequently used in process mining projects in

industry. Moreover, some of the ideas and algorithms have been incorporated in commercial BPM tools like BPMone (Pallas Athena), Futura Reflect (Futura Process Intelligence), ARIS PPM (IDS Scheer), etc.

Figure1

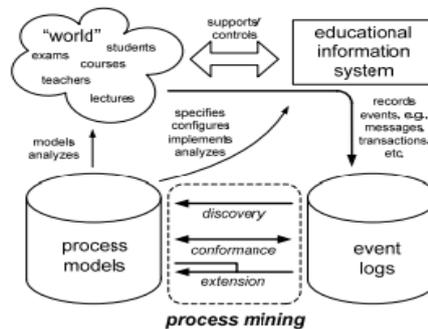


Figure 1. The process mining spectrum supported by ProM

The first type of process mining is discovery; it takes an event log and produces a model without using any a-priori information. Process discovery is the important process mining technique. A process discovery algorithm is to construct a process model, which reflects the behavior that has been observed in the event log. Different process modeling languages can be used to capture the causal relationships of the steps, or activities, in the process.

The second type of process mining is conformance. In this technique, an existing process model is compared with an event log of the same process. Conformance checking can be used to check if reality, as recorded in the log, conforms to the model and vice versa. Typically, four quality dimensions for comparing model and log are considered: (a) fitness, (b) precision, (c) generalization and (d) Structure.

The Third type of process mining is enhancement. This technique is used to give the idea to extend or improve an existing process model using information about the actual process recorded in some event log. It aims at changing or extending the a-priori model. A non-fitting process model can be corrected using the diagnostics provided by the alignment of model and log.

III. Process Mining Perspectives

Process mining is one such tool to extract non trivial and useful process information from event logs. Process mining is a technique that takes event log from a business process as its input and tries to generate a model for the business process that generated the logs For mining business process there are different perspectives: 1) The process perspectives .2) The Organizational perspectives 3) The case perspectives.

The process perspectives (HOW?): It focuses on the control flow (i.e.,) the ordering of activities, expressed in terms of, a petri net or Event-driven Process Chain (EPC)

The Organizational perspectives (Who?): It focuses on the originator (i.e.) which performers are involved in performing the activities and how they are related. For mining the organizational perspectives, the emphasis can be on the roles or the social network or on the utilization of performers or execution frequencies.

The case perspectives (What?): It can be characterized by their path in the process or by the originators working on a case and can also be characterized by the values of the corresponding data elements. The case perspective looks at the case as a whole and tries to establish relations between the various properties of a case.

Resource Perspective: The resource perspective looks at the set of people involved in the process, and their relationships. The Social Network Miner for example can generate the social network of the organization, which may highlight different relationships between the persons involved in the process, such as handover of work, subcontracting and others. The Organizational Miner also addresses the resource perspective, attempting to cluster resources, which perform similar tasks into roles.

Performance Perspective: Mining algorithms addressing the performance perspective mainly make use of the timestamp attribute of events. From the combination of a (mined or predefined) process model and a timed event log, they can give detailed information about performance deficiencies, and their location in the process model.

Information Perspective: The Activity Miner can derive high-level activities from a log by clustering similar sets of low-level events that are found to occur together frequently. These high-level clusters, or patterns, are helpful for unveiling hidden dependencies between documents, or for a re-structuring of the document repository layout.

IV. The Integration Of Process Mining In Higher Education Topics

Process mining is an innovative field of research and study, which can be implemented in education with and its perspectives can be applied and put into practice in the academic world.

A main concern of each institution of higher education is to predict the paths of students and alumni (Luan, 2004). They would like to identify which students will join particular course programs, and which students will require assistance in order to graduate. At the same time institutions want to learn whether some students more likely to transfer than others, and what groups of alumni are most likely to offer pledges. In addition to this challenge, traditional issues such as enrollment management and time-to-degree continue to motivate higher education institutions to search for better solutions.

Various studies are based on students' present "knowledge luggage", detecting with realistic accuracy what and how much the students know, in order to deduce specific learning gaps. This set of information could be obtained during an ongoing learning assessment process that makes possible to specify, with reasonable precision, which subject the student is better suited to learn at that moment, and requires automatic or semiautomatic procedures for treatment and analysis for acquisition of new knowledge. Pimentel and Omar (2005) presented a model for organizing and measuring knowledge upgrade in systems of education and learning with the support of data mining tools.

Vandamme and collaborators (2007) attempted to classify, as early in the academic year as possible, students into three groups: the 'low-risk' students, with a high probability of succeeding; the 'medium-risk' students, who may succeed thanks to the measures taken by the university; and the 'high-risk' students, with a high probability of failing (or dropping out). They present the results of their application of discriminate analysis.

Since universities courses progressively oblige students to use online tools in their studies, there are numerous prospects to mine the resultant large quantity of student learning data for hidden valuable information. Several approaches in scientific literature try to classify students in order to predict their final grade based on features extracted from logged data in educational web-based systems (Minaei-Bidgoli & Punch, 2003).

Intelligent tutoring systems' capability to register its exchanges with students is an important challenge and an opportunity (Mostow et al., 2005). A central matter in mining event logs from an intelligent tutoring system is "What happened when...?". In comparison to individual observation of live or videotaped tutoring, logs can be more far-reaching in the number of students, more comprehensive in the number of sessions, and exquisite in details, avoiding thus observer effects, costing less to obtain, and easier to analyze. Mostow and collaborators describe (2005) an educational data mining tool to support such case analysis by exploiting three simple but powerful ideas: first, a student, computer, and time interval suffice to specify an event; second, a containment relation between time intervals defines a hierarchical structure of tutorial interactions; third, the first two ideas make it possible to implement a generic but flexible tool for mining tutor data with minimal dependency on tutor-specific details.

The Academic Intelligence Management includes all higher education institution's processes utilized to acquire, generate and spread knowledge in order to accomplish its objectives and Strategies based on the latest ICT (Information and Communication Technologies) and collaborative practices. It is tied to organizational goals such as improved performance, competitive advantage, innovation, research and development, which derive from technologies and applications providing historical, present, and predictive analysis of all academic activities. The ways in which information and knowledge are represented and delivered to the university managers are in a continuous transformation due to the involvement of the information and communication technologies in all the higher education processes.

The higher education institutions represent dissimilar and complex environments, which involve links to communication, and collaboration among its various departments and the society. The University is progressively regarded as a collaborative organization whose mission is to foster knowledge creation and knowledge diffusion among communities of students, scholars and researchers (Rodrigues & Barrulas, 2003). In the case of higher education institutions, the designing circuits of a managerial system (Rusu & Bresfelean, 2006) aiming an academic intelligent/intelligence management must be closely tied to:

- Educational process activities: Structure design and curricula content must be token of labor market requests and institution's capability; quality and freshness of courses information; adequate teaching/learning and evaluation methods; appropriate performance of educational processes etc.
- Scientific research activities: Thematic originality and opportunity; consistency of results; scientific probability; ethic experimenting; ways of utilizing the results etc.
- Internal organization: Authority and responsibility delegation; transparency and efficiency in the utilization of human and material resources; equity and performance encouragement in personnel promotion; continuous personnel training etc.
- External relations with local, national and international community; relations with other educational institutions and companies from different activity sectors; alumni etc

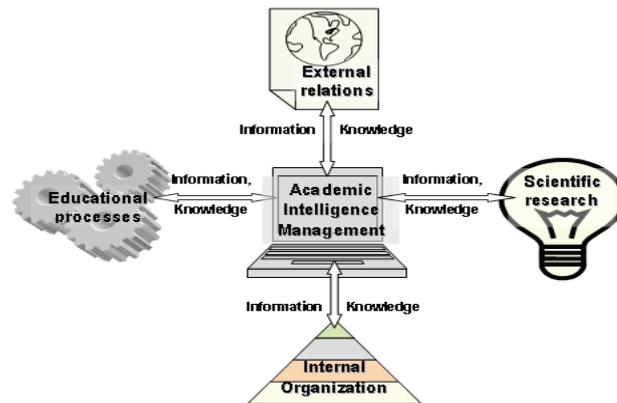


Figure. 2. Academic Intelligence Management, the hub of higher education activities

The studies will continue with deeper mining in the process of educational topics, such as performance in scientific research, correlations between the students' knowledge and the competences demanded on the labor market, academic failure, to perceive what and how much the students know, to realize learning gaps, and also improve teaching methods and educational management processes.

V. Conclusion

Process mining technique even though largely applied in the business management activities, this paper suggests the higher education environment also includes lot of activities, mentioned in the literature, can also utilize the process mining techniques to improve the learners performance. The academic and administrative work in all the departments, across the departments and collaboration with other systems in the society can also be analyzed using process mining tools. Further research work of this paper continues to analyze the event logs in all the perspectives of process mining techniques in educational domain.

References

- [1] Minaei-Bidgoli B., Punch W.F., Using Genetic Algorithms for Data Mining Optimization in an Educational Web-based System, GECCO 2003 Conference, Springer-Verlag, Vol2, Chicago, USA; July 2003. pp.2252-2263.
- [2] Mostow J., Beck J., Cen H., Cuneo A., Gouvea E., and Heiner C., An educational data mining tool to browse tutor-student interactions: Time will tell! Proceedings of the Workshop on Educational Data Mining, Pittsburgh, USA; 2005. pp.15-22.
- [3] Pimentel E.P., Omar N., Towards a model for organizing and measuring knowledge upgrade in education with data mining, The 2005 IEEE International Conference on Information Reuse and Integration, Las Vegas, USA; August 15-17, 2005
- [4] Rodrigues, J.P.C., Barrulas, M. J. (2003), Towards Web-Based Information and Knowledge Management in Higher Education Institutions, Lecture Notes in Computer Science, Volume 2720, Sep 2003, pp. 188-197
- [5] Rusu, L., Breşfelean, V.P., Management prototype for universities. Annals of the Tiberiu Popoviciu Seminar, Supplement: International Workshop in Collaborative systems, Volume 4, 2006, Mediamira Science Publisher, Cluj-Napoca, Romania, pp. 287-295
- [6] Vanda me J.P., Meskens N., Superby J.F., Predicting Academic Performance by Data Mining Methods, Education Economics, Volume 15, Issue 4 December 2007 , pp.405 – 419
- [7] Van der Aalst, W.M.P., Weijters, A.J.M.M., Maruster, L. Workflow Mining: Discovering Process Models from Event Logs. *IEEE Transactions on Knowledge and Data. Engineering* 16(9), p.1128–1142, 2004.
- [8] Vasile Paul Bresfelean Data Mining Applications in Higher Education and Academic intelligence Management 27. June 2008 Online at <http://mpr.ub.uni-muenchen.de/21235/> MPR Paper No. 21235, posted 13. March 2010 10:55 UTC
- [9] Rozinat, A., van der Aalst, W.M.P. Conformance Checking of Processes Based on Monitoring Real Behavior. *Information Systems* 33(1), p. 64-95.
- [10] Vladimir Rubin, Christian W. Günther1, Wil M.P. van der Aalst, Ekkart Kindler, Boudewijn F. van Dongen, and Wilhelm Schöfer Process Mining framework for Software Processes
- [11] Luan Jing, Data Mining Applications in Higher Education, SPSS Exec.Report,2004, http://www.spss.com/home_page/wp2.htm
- [12] R.S Mans;M.H.Schneberg;M.Song;W.M.P.Van der Aalst, Eindhoven University of Technology, P.O.Box 513,NL-5600 MB,Eindhoven,The Netherlands,Process mining in healthcare, A case study.
- [13] Mykola Pechenizkiy, Nikola Trčka, Ekaterina Vasilyeva, Wil van der Aalst, Paul De Bra{m.pechenizkiy, e.vasilyeva, n.trcka, w.m.p.v.d.aalst}@tue.nl, debra@win.tue.nl Department of Computer Science, Eindhoven University of Technology, the Netherlands, Process Mining Online Assessment Data.