

AperTO - Archivio Istituzionale Open Access dell'Università di Torino

Enhancing E-learning effectiveness: a process mining approach for short-term tutorials

This is the author's manuscript

Original Citation:

Availability:

This version is available <http://hdl.handle.net/2318/2016072> since 2024-09-25T08:34:37Z

Published version:

DOI:10.1007/s10844-024-00874-9

Terms of use:

Open Access

Anyone can freely access the full text of works made available as "Open Access". Works made available under a Creative Commons license can be used according to the terms and conditions of said license. Use of all other works requires consent of the right holder (author or publisher) if not exempted from copyright protection by the applicable law.

(Article begins on next page)

Enhancing E-Learning Effectiveness: A Process Mining Approach for Short-Term Tutorials

Roberto Nai^{1*}, Emilio Sulis¹ and Laura Genga²

^{1*}Computer Science Department, University of Turin, Corso Svizzera, 185, Turin, 10149, Italy, IT.

²Industrial Engineering and Innovation Sciences, Eindhoven University of Technology (TU/e), De Zaale, Eindhoven, 5600, The Netherlands, NL.

*Corresponding author(s). E-mail(s): roberto.nai@unito.it;
Contributing authors: emilio.sulis@unito.it; l.genga@tue.nl;

Abstract

The rise of e-learning systems has revolutionized education, enabling the collection of valuable students' activity data for continuous improvement. While existing studies have predominantly focused on prolonged learning paths, short-term tutorials offer a flexible and efficient alternative that is recently gaining increasing popularity. This article presents a methodology for investigating e-learning systems for short-term tutorials leveraging user behavior tracking and process mining techniques. A case study involving a web-based tutorial with approximately one hour of learning explores the learning processes of 250 students in Italy. The study analyzes learning outcomes and investigates the impact of different learning paths on student progress. The research questions concern i) the extraction of activity flows in short-term tutorials; ii) the prediction of outcomes in the early stages of short-term learning process. The proposed approach provides descriptive insights into the learning process which can also be used to offer prescriptive guidance.

Keywords: Tutorial design, User behaviour tracking, Process Mining, Predictive Process Monitoring, Learning outcomes

1 Introduction

E-learning systems are digital platforms facilitating online education and training, delivering educational content and interactive experiences via the Internet. These systems support remote learning, often incorporating multimedia elements, assessments, and collaboration tools to enhance the educational experience. Moreover, they can record data on the activities carried out by students during their learning process.

Several previous studies have investigated the adoption of automated techniques to extract knowledge from information system (IS) data and improve learning knowledge (Kabudi et al (2021)). For instance, it has been demonstrated how helpful information can be extracted to identify learning patterns and provide recommendations on what to study next (Feng and Fan (2024)). This body of research has primarily focused on learning activities having a relatively long duration, typically several months or years. However, short tutorials on specific topics are becoming increasingly popular via the Web. These short tutorials make it possible to introduce complex topics in learning units of only a few hours. These flexible tools can adapt to individual students' schedules and learning preferences.

Despite their increasing popularity, their analysis and assessment have received little attention in the literature. In particular, to the best of our knowledge, no previous study has investigated the feasibility of applying evidence-based educational data mining techniques to data generated from these tutorials. When considering complete courses or, anyway, long-term learning activities, the analysis can usually rely on a large volume of data, encompassing multiple modules, assignments, and exams, which support the extraction of detailed and nuanced insights into student behavior and learning patterns. When analyzing short tutorials, the data volume is usually much smaller, often limited to a single session or a few interactions. This can make it easier to analyze but may provide less comprehensive insight. This paper aims to fill this gap. More precisely, we aim to investigate the feasibility of leveraging EDM techniques to provide valuable insights into the learning processes underlying short-tutorials. We argue that such insights can be used to refine specific teaching methods, improve individual tutorial content, and provide rapid feedback to both students and educators. In particular, we propose a methodology for collecting data in an IS by tracking user behavior to evaluate short tutorials. The methodology we propose can collect timed events that can then be analyzed through techniques of the Process Mining (PM) discipline, namely process discovery, variant analysis, and predictive process monitoring. Such process-aware analysis techniques have already shown their benefits in a variety of different learning processes and analysis objectives in the Educational Process Mining discipline (Bogarín et al (2018)). To showcase the capabilities of our methodology, a case study was conducted through the creation and administration of an instructional tutorial employing Web-based technologies capable of collecting data during the learning process for a short period of time (approximately one to two hours). We address the learning process of about 250 Italian students in an introductory programming course.

The first research objective involves the automated extraction of useful information in short-term learning process. Special attention has been paid to analyse learning outcomes to identify success factors based on the educational journey. In particular, we

aim to uncover students’ activity flows, considering both aggregated process indicators and process variants. A second research objective concerns the ability to predict the tutorial’s outcome starting from the initial steps.

Our methodology concretely supports the above-mentioned research objectives. First, it allows us to mine students’ learning processes and uncover potential relations between different variants and learning outcomes. The corresponding research question concerns extracting students’ activity flows in a short-term learning process and assessing whether there is any relation to the outcome (RQ1). Second, focusing mostly on prediction, we analyse the data from our case study to estimate the relationship between behavior in performing the tutorial and learning. The research question concerns the ability to predict students’ outcomes at different stages of the learning process in short-term tutorials (RQ2).

To validate the reliability of the obtained insights, we considered the teachers involved in the administration of the tutorial as domain experts. We add the discussion of the results obtained and the teachers’ comments in the case study directly to the different parts of the results analysis.

In the remainder of the paper, we introduce the case study in Section 2 and the methodology in Section 3. We finally examine the results in Section 4, providing a discussion in Section 5, while we review and compare the related work in Section 6. Finally, Section 7 concludes the paper.

2 Case Study

We discuss a case study based on a web tutorial to track students’ behavior during their learning process. We administered the tutorial to two groups of undergraduate students enrolled in the second year of their degree programs in management and economy. The educational context is that of an introductory computer science course, with a class of students homogeneous in age group and fairly evenly distributed by gender. The students are prompted by the teachers administering the tutorial to take the course individually and, in general, are left free to possibly interact, with no control over their individual behavior. This work expands on the idea presented in a previous paper that illustrated the framework and early results of the application to 70 students while also proposing the possibility of a qualitative follow-up with ethnographic research (Nai et al (2023)).

Tutorial web

Our tutorial consists of 10 web pages on topics related to learning Python programming. Each page is a self-contained introductory lesson that can be performed without prior knowledge. A multiple-choice question from each of the 10 pages tests whether students follow and learn the topics covered. In the end, the sum of the correct answers suggests whether they are learning well or poorly (see Section 3.3).

To investigate the learning sequence of the subject taught, we propose three different paths. Figure 1 represents the sequence of lessons proposed in our tutorial according to the Business Process Model and Notation (BPMN) (von Rosing et al (2015)). The following four lessons out of ten are the same for all students: the final one

(FUNCTIONS), as well as the first three (INTRODUCTION, FIRST PROGRAM, and VARIABLES). After the third lesson about variables, the student can choose one of three learning paths in which the order of the three topics presented changes.

In particular, the following six lessons are included in three topics (i.e., three topics of two lessons each):

- **DATA TYPES:** a lesson to introduce different data types (TYPES) and a lesson on the conversion between different data types (CONV);
- **DATA STRUCTURES:** a lesson on the collection of similar data items (LISTS) and a lesson about unordered collection of keys and values (DICTS);
- **CONTROL STRUCTURES:** a lesson to execute a block of code according to specific conditions (IF_ELSE) and a lesson to repeat a block of code until a specific condition (FOR).

To the best of our knowledge, the *Track1* (DATA TYPES, DATA STRUCTURES, CONTROL STRUCTURES) is the most typical order according to contemporary computer science manuals. Nevertheless, in our tutorial, learners can follow one of two other tracks with the same contents but in a different order: *Track2* (DATA STRUCTURES, DATA TYPES, CONTROL STRUCTURES) or *Track3* (CONTROL STRUCTURES, DATA TYPES, DATA STRUCTURES). For this specific analysis objective, *Track1* is our benchmark for comparison. *Track3* is the furthest from the ideal path, according to the domain experts involved.

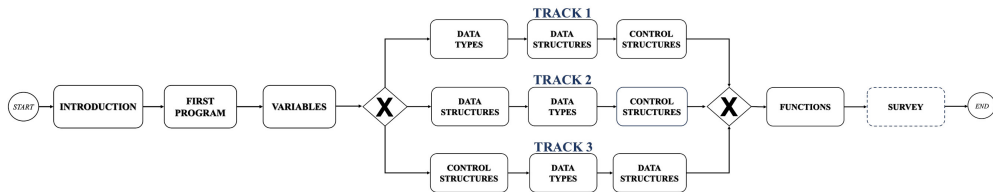


Fig. 1 The three different learning paths; on the top, the standard learning flow (*Track1*). Full size image available at <https://bit.ly/3Tm6GJD>

3 Methodology

The methodology adopted in the present work includes several phases, as summarized in Figure 2. The first phase concerns the *design* stage, which involves defining the tutorial’s content and quizzes. In the second phase, the *data collection* involved the development and administration of a web tutorial that incorporated the previously defined content and quizzes. The student activity monitoring system ensures that all interactions with the tutorial are recorded accurately. In the *event log construction* phase we converted the collected tracking data into an event log. This process involved structuring the raw data into a format suitable for analysis, allowing us to trace each student’s learning through the tutorial. In the *variant analysis* phase, we analysed the event log to identify how students interacted with the tutorial and identify patterns

that might influence learning outcomes. Finally, in the *outcome prediction* phase, we applied predictive analytics to the event log data. This involved using the insights gained from the variant analysis to predict future outcomes, such as student performance and areas where students might struggle. The predictions made in this stage aimed to inform educators and improve the tutorial’s effectiveness.

Each phase - design, data collection, event log construction, variant analysis, and outcome prediction - was integral to the comprehensive evaluation and enhancement of the web tutorial.

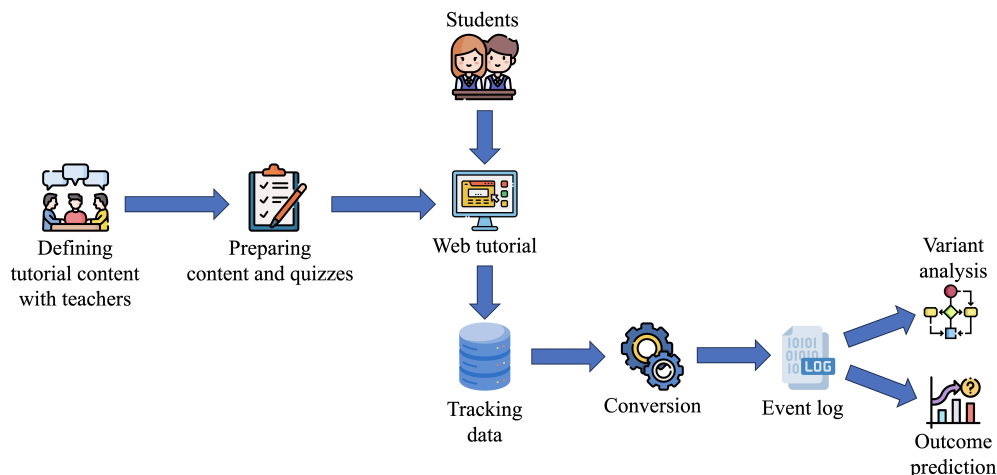


Fig. 2 Summary of our case study’s phases: definition of the tutorial content, web portal for the administration of the content, activity tracking, convert tracking into an event log, variant analysis, and prediction on the event log. Full size image available at <https://bit.ly/3Tm6GJD>

3.1 Web-tracking and technologies

Web behavior tracking. Each page/lesson is divided into 3-4 paragraphs describing the topics. While browsing the tutorial, the following events were tracked within the web pages: PAGE-IN (enter the page), PAGE-OUT (exit the page to access the following lesson or quiz), MOUSE-IN (the mouse pointer moves onto a new paragraph element), MOUSE-OUT (the mouse pointer moves out of a paragraph element), CLICK (a student clicks on a paragraph element), DBCLICK (a student double-clicks on a paragraph element). Each page also tracks movement between paragraphs (numbered 0 to 3). E.g., INTRO_MOUSE-IN_2 means that the student has entered with the mouse in the second paragraph of the introductory lesson. In addition to the events, the following data are recorded: the session-id (to trace the activities back to a specific browser session), the name of the page (INTRODUCTION, FIRST PROGRAM, VARIABLES, etc.) on which the events occurred, the result of the quiz on that page, the *Track* on which the events occurred (1, 2 or 3), and optional data provided by the students at the time of the final survey (e.g., tutorial evaluation).

Technologies. The tutorial described in Section 2 has been implemented by adopting the following technologies to track behavior in web pages during the course of the tutorial. The *front-end* (the graphical user interface) and the contents of each tutorial page are written in HyperText Markup Language (HTML), using the open-source frameworks Bootstrap¹ for the layout and jQuery² for event tracking, based on (Ermakova et al (2018); Bujlow et al (2017)). The *back-end* programming language for track functionality called up via jQuery is PHP³ while the Database Management System (DBMS) for storing and retrieving tracking data is MySQL.⁴ For web tracking, the following steps were performed: a session-id was created when a student first accessed the web tutorial; all sections (paragraphs) of the tutorial pages were labelled in HTML and, based on the student’s interaction with the pages, a jQuery function performed an asynchronous call to the server to track the interaction, saving the DBMS table using the session-id as the key entry. The web tutorial is available online.⁵

3.2 Event log construction

The starting point for PM research is an *event log*, representing an extraction from the IS on the execution of activities in a process (van der Aalst (2016)). An event log includes a set of *traces*, whereas each trace stores a sequence of *events*, each representing the execution of an *activity* occurred at a given *timestamp* during a process, possibly together with some additional data (e.g., the resource who performed the activity) (van der Aalst (2016)). Every trace is identified using a so-called *Case_ID*, which is the session-id assigned automatically by the web browser to a student when navigating with the browser within the tutorial. The fields Activity and Timestamp are the data traced by jQuery during the tutorial execution. Additional information, such as data provided by students, is added for each track (e.g., track choice). To focus on the learning flow, we studied four kinds of activities: the entrance and exit on a web page (PAGE-IN, PAGE-OUT) and the entrance and exit in each paragraph of the page (MOUSE-IN, MOUSE-OUT). The number of CLICKs and DBCLICKs were instead used as trace features.

Different groups of traces in the event log are called *process variants* (*variants*) since they represent alternative ways to execute the same process (i.e., users may perform activities in a different order before the end). After exploring the dataset, some traces configured as outliers can be removed if they appear to be wrong or not harmonious, such as in the case of processes that are too short according to domain experts (e.g., students who perhaps only opened the initial pages without proceeding with the tutorial). To focus on the most significant cases, we consider students having completed the tutorial in between 5 minutes and 2 hours and a half. The tracking data model, the scripts used in the current work, and the event logs are publicly available⁶.

¹<https://getbootstrap.com>

²<https://jquery.com>

³<https://www.php.net>

⁴<https://www.mysql.com>

⁵<http://webtutorial.altervista.org/python>

⁶Repository of material associated with this article: <https://bit.ly/3Tm6GJD>

3.3 Outcome analysis

A relevant analysis in the present work concerns the distinction between students who performed well and poorly. To have an indicator of whether students have understood the tutorial content, we rely on the 10 answers given to the quizzes between each page, counting one point for each correct answer. The distribution of the results can be divided into two parts through the median class as a threshold; in our case study, the median value obtained from the quiz results used as a threshold to distinguish the two parts is 0.7. As discrete data/classes, we obtained two groups of almost equal size. Finally, in the current proof-of-concept, we opted to define processes with *negative outcome* (OUT-NEG) as all cases below the threshold, as well as processes with *positive outcome* (OUT-POS) above the threshold. The value of each student's outcome has been saved in the event log.

3.4 Process mining techniques

Variant analysis (RQ1)

A first exploration involves the analysis of event logs and diagrams obtained from process discovery. We inspect both the complete log and the individual processes of interest according to our RQ1. First, we investigated the students' performance concerning the corresponding learning outcomes (Section 4.2). In particular, we intend to analyse the log according to several dimensions to identify interesting behaviors. The variants will be considered in relation to the overall completion time of the tutorial, the time spent on each page, and the student's movements between paragraphs and pages of the tutorial. We also proceeded with an *automated* comparison aimed at quantifying the existing differences. More precisely, in this work, we apply the approach proposed by Bolt et al (2018), which takes into account both *behavioral* and *context* process similarity. The first one considers how activities are executed in the compared executions. The second one takes into account the context in which the executions occur, defined using the data attributes stored in the event log. The approach takes as input two event logs corresponding to the set of executions to compare. Then, it computes the differences among them in terms of behavior or context and builds a transition system representing the behavior of both variants, where states or edges showing relevant differences between the two variants are annotated accordingly. Note that these annotations are visualized using different colours and thicknesses of the transition system elements.

Second, we compared the three discovered processes about *Track1*, *Track2*, and *Track3* to investigate any differences of interest (Section 4.3). The first analysis concerns tracking timing and outcome. Second, we focus on the times between individual lessons in the three tracks. Finally, we examine the backward jumps between paragraphs and pages in each track (intended as the action of returning to a previous section or page of the tutorial). We intend to examine this behavior as it may indicate a desire for better learning or distraction, according to the domain experts.

To analyse the event log, we used academically licensed DISCO from Fluxicon⁷, as well as ProM⁸ to perform the automated variant analysis (*process-comparator* module).

Predictive Process Monitoring (RQ2)

Predictive Process Monitoring (PPM) (Maggi et al (2014)) is a branch of PM research that aims to predict the future development of ongoing process cases given their uncompleted traces. According to our RQ2, we aim to predict students' performance based on the learning process taken by the students in earlier stages (an *outcome-based* prediction). Figure 5 summarises the phases of our PPM exploration. First, from the complete event log, we refer to the sequences of events recorded up to a certain point in time during the execution of a process. These partial event sequences are called *prefixes* to be used for predicting the future behavior of the process. In the training phase on machine learning models, *prefixes* extracted from the traces of the event log (Di Francescomarino and Ghidini (2022)) become vectors according to different encoding techniques. In our research, we used *Index Encoding* (IE), *Boolean Encoding* (BE) and *Frequency Encoding* (FE) (Di Francescomarino and Ghidini (2022)) methods to verify which one leads to better results with the available data. In particular, in IE, each feature corresponds to a position order in the sequence, and the possible values for each feature are the event classes; BE represents a sequence through a feature where an event is indicated by 1 if it occurred in the prefix, 0 otherwise; FE represents the control flow in a case with the frequency of each event class in the case. Figure 3 describes an example of the three encodings; the IE (Figure 3.a) includes the sequence of events that occurred for each Case ID (e.g., for Case 'ID01', the first event occurred IF_ELSE_PageIN_0), and the third is IF_ELSE_MouseIN_1). BE (Figure 3.b) assigns 1 for events that occurred and 0 for those that did not occur for each Case ID (e.g., for Case 'ID01', the event IF_ELSE_PageIN_0 occurred while IF_ELSE_MouseOUT_1 did not). Finally, FE (Figure 3.c) includes the frequency with which the events occurred (e.g., for case 'ID01', the event IF_ELSE_PageIN_0 occurred 3 times).


Finally, supervised experiments are applied to these trace representations to obtain a predictive model. Such a model can then be applied to new partial traces. At runtime, predictions are made on incomplete traces. Since our research aims to make predictions as early as possible, we focused on the subset of the prefix log with the initial part of the process, i.e. a length of 40 (which corresponds to the first page/lesson of the tutorial), 80 (which corresponds to the second page/lesson of the tutorial), or 160 (which corresponds to the third page/lesson of the tutorial). We trained two single classifiers: Random Forest (RF) and eXtreme Gradient Boosting (XGB). The traces in input to classifiers are zero-padded to have a fixed length.

Figure 4 graphically shows a complete trace of length n (Figure 4.a) as well as the trace prefixes of length 1 (Figure 4.b), length 2 (Figure 4.c), and length 3 (Figure 4.d) with zero-padding.


Outcome prediction In terms of technology, we used the open-source toolkit Nirdizati (Rizzi et al (2022)), which supports the various phases of the PPM just

⁷<https://fluxicon.com/disco>


⁸<https://promtools.org>

Case ID 	Event 1	Event 2	Event 3	...	Event n
ID01	IF_ELSE_PageIN_0	IF_ELSE_MouseIN_1	IF_ELSE_MouseOUT_1
ID02	IF_ELSE_PageIN_0	IF_ELSE_MouseIN_2	IF_ELSE_MouseOUT_3
...

(a) – Index Encoding (IE)

Case ID 	IF_ELSE_PageIN_0	IF_ELSE_MouseIN_1	IF_ELSE_MouseOUT_1	...	Event n
ID01	1	1	0	...	0
ID02	1	1	0	...	1
...

(b) – Boolean Encoding (BE)

Case ID 	IF_ELSE_PageIN_0	IF_ELSE_MouseIN_1	IF_ELSE_MouseOUT_1	...	Event n
ID01	3	4	5	...	0
ID02	2	1	6	...	1
...

(c) – Frequency Encoding (FE)

Fig. 3 Examples of prefixes encoded with simple-index (a), boolean (b) and frequency (c). Full size image available at <https://bit.ly/3Tm6GJD>

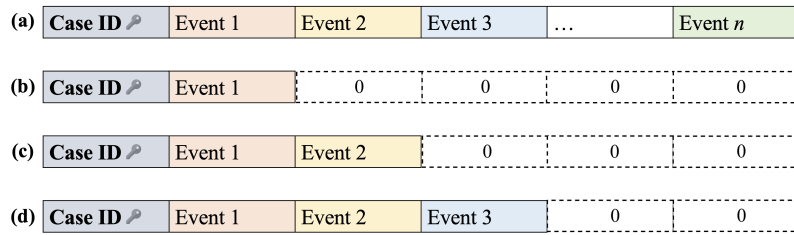


Fig. 4 Examples of prefixes starting from a complete trace (a) and related trace prefix of length 1 (b), length 2 (c), and length 3 (d), all with zero-padding to have the same length. Full size image available at <https://bit.ly/3Tm6GJD>

described.⁹ Table 1 summarises the trace features used as input for the prediction models. The output is the binary classification between *positive outcome* (OUT-POS) or *negative outcome* (OUT-NEG).

Table 1 Trace features used as input for the prediction models

Trace feature	Description
event	Encoded version of the event
trace_total_time	Total trace time
track	Tutorial track {1, 2, 3}
student_age	Age of the student
student_gender	Gender of the student {M, F, other}
click_num	Number of clicks
dclick_num	Number of double clicks
page_jump_num	Number of backward jump per page
para_jump_num	Number of backward jump per page paragraph

⁹<http://research.nirdizati.org>

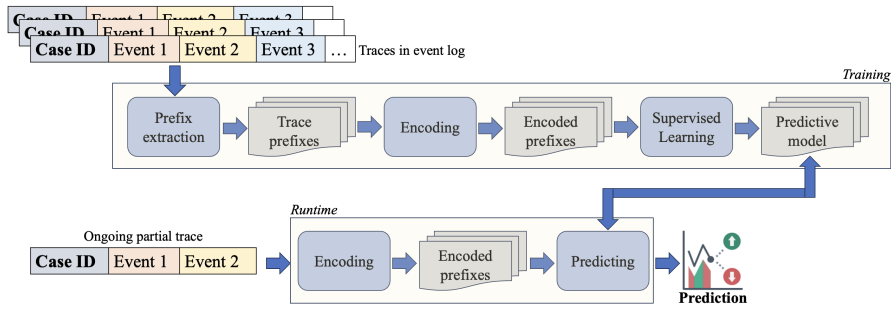


Fig. 5 Overview of our PPM exploration based on machine learning models. Full size image available at <https://bit.ly/3Tm6GJD>

The prediction results are evaluated with *K-fold cross-validation* and *F1-Score*, i.e. the harmonic mean between *recall* and *precision* (Géron (2022)), and the *Area Under the Curve* (AUC) (Fawcett (2006)). The F1-Score metric is a unique measure of models’ prediction performance with an imbalanced dataset (Buckland and Gey (1994)), while the AUC metric is calculated by assessing a classification model’s ability to distinguish between classes (Fawcett (2006)). The hyperparameters optimisation used by Nirdizati is Hyperopt (Bergstra et al (2015)).

The computations were carried out by an ARM architecture-based chip with 3.2 GHz speed (10-Core CPU / 24-core GPU) and 32 GB of RAM.

4 Results

4.1 Event log analysis

The complete dataset includes sessions from the tutorial administrations. Table 2 shows the main statistics of the event log; most students concluded the tutorial with a median duration of 37.8 minutes as well as an average duration of 41.3 minutes. The standard deviation (STD) of 28.9 minutes is quite relevant; in fact, there is considerable variability. In the most extreme cases, some students completed the tutorial very quickly (5 minutes) while, on the contrary, a few students needed 2 hours to complete it.

Table 3 shows a snapshot of the resulting event log, with the three main properties in the event log (Case ID, Activity, Timestamp) and an example of the other features we added as attributes of the traces, namely the type of track that the student travelled. According to the final survey, the 82% of students expressed high appreciation for the tutorial. This seems relevant both to ensure the effectiveness of the proposed approach and to proceed with the examination of the results.

4.2 Learning processes and outcome analysis

Analysis of the learning process’ timing

In terms of time analysis, we focus on the overall duration of the learning process and the time spent on individual tutorial pages. The duration of the learning processes

Table 2 Main statistics on the event log obtained from the tutorial: the first line shows the statistics of all cases, the second line those of cases with a positive outcome (OUT-POS), the third line shows those with a negative outcome (OUT-NEG). Times are expressed in hours (h), minutes (m), and seconds (s)

Cases	Total	Median duration	Mean duration	STD	Min duration	Max duration	Backward jumps (avg)
Complete	242	37.8 m	41.3 m	28.9 m	5 m, 8 s	2 h, 19 m	1.64
OUT-POS	130	42.7 m	45.8 m	32.8 m	5 m, 8 s	2 h, 2 m	1.79
OUT-NEG	112	32 m	36.1 m	21.7 m	5 m, 51 s	2 h, 19 m	1.56

Table 3 A sample example of the event log including the activities of a single student identified with Case ID ‘ID01’, navigating to the IF-ELSE, FOR, TYPES, and LISTS web pages in *Track3* learning path (‘Track’ and ‘Quiz’ are trace features)

Case ID	Activity	Timestamp	Track	Quiz
ID01	IF_ELSE_PageIN_0	2023-03-29 17:38:50	3	0.8
	IF_ELSE_MouseIN_1	2023-03-29 17:39:13		
	IF_ELSE_MouseOUT_1	2023-03-29 17:39:23		
	IF_ELSE-Q_PageIN_0	2023-03-29 17:40:02		
	IF_ELSE-Q_PageOUT_0	2023-03-29 17:40:15		
	FOR_PageIN_0	2023-03-29 17:40:16		
	FOR_MouseIN_1	2023-03-29 17:40:20		
	FOR_MouseOUT_1	2023-03-29 17:40:24		
	FOR_MouseIN_2	2023-03-29 17:40:31		
	TYPES_PageIN_0	2023-03-29 17:46:44		
	TYPES_MouseIN_2	2023-03-29 17:46:57		
	TYPES_MouseOUT_2	2023-03-29 17:47:01		
	LISTS_PageIN_0	2023-03-29 17:55:41		
	LISTS_MouseIN_1	2023-03-29 17:55:42		
	LISTS_MouseOUT_1	2023-03-29 17:55:48		
LISTS_MouseIN_2	2023-03-29 17:57:56			

(median and average duration) clearly indicates that students with positive outcome took longer. As summarized in Table 2, the median duration of the tutorial is about 46 minutes for students with positive outcomes, while students while students with negative outcomes took a median time of 32 minutes. Concerning students with poor performance spending less time consulting computer tutorials could be attributed to several factors. A hypothesis is that these students may lack the necessary foundational knowledge to engage effectively with the tutorial content. As a result, they may rush through the material without fully understanding it, leading to lower performance outcomes (Sweller (1994)).

The behavior on individual pages of the students with positive/negative outcome is also analyzed. We note how the top-performing students were slower for each of the ten pages, and we identify a significantly longer median duration than those who performed poorly. As Table 4 highlights, times on pages are always broadly higher for the group that will get a successful outcome. The stay on the pages can often be more than twice as long. Interestingly, this behavior appears already in the first pages, suggesting a student’s attitude that can thus be intercepted as early as the first part of the tutorial execution.

Table 4 The duration (in seconds) on individual pages for the group of cases with positive outcome (OUT-POS) and negative outcome (OUT-NEG)

	INTRO	PROG	VARs	IFELSE	FOR	TYPES	CONV	LIST	DICT	FUNCT
OUT-NEG	66	79,5	71	79	59	126	81	61	58,5	28
OUT-POS	120	120	138	138	116	294	150	150	82	58

The movements between pages or paragraphs

By observing the jumps between different pages or activities (i.e. paragraphs) during the course of the tutorial (Table 2), we can observe a meaningful difference in students' behavior in carrying out the tutorial. We computed the average number of backward jumps in relation to the learning outcome. The group of students with positive outcome appears to go back more frequently (1.79 jumps backward on average), with less linear behavior, than those with negative outcome (1.56 jumps backward on average). A behavior perhaps aimed at improving contents understanding, corresponding to a more reflective attitude.

Together with the previous observation about timing, the result seems to indicate that students with positive outcome focus more carefully on the content and return to topics already covered, while those with negative outcome proceed quickly towards the next paragraph, without going back very often and making sure they have understood the tutorial content.

Automated variant analysis results

Finally, we perform a statistical comparison of subgroups' traces with positive and negative outcomes (as mentioned in Section 3.3, we exploit *Process Comparator* plugin in ProM tool). Such a comparison allows to identify which parts of the tutorial appear relatively more significant. Figure 6 reports the obtained results, whereas the darker the color tone, the stronger the statistical relevance of the difference between activities.

To provide an idea of this type of analysis, we describe three cases that are of interest. First, we focus on the frequency of activities. In Figure 6.a, the central paragraphs in the pages concerning conversions (CONV) appear relatively more frequent among students with positive outcome.

Second, in Figure 6.b, log comparison indicates that there are statistically significant differences in terms of duration for performing the activities in the section LISTS. Cases with positive outcomes spent more time, compared to negative ones, on the paragraphs related to LISTS learning.

Third, regarding the differences between activities with respect to the corresponding remaining times, the diagram in Figure 6.c shows that there are differences in INTRO and PROG sections. Being the initial activities of the tutorial, this observation confirms what we had already found in the analysis about timing, namely that students with positive outcome take longer from the tutorial beginning to finish the activities.

These results illustrate the possibilities offered by this type of automatic analysis. Overall, these suggestions may indicate the parts of the tutorial to focus on to propose possible improvements.

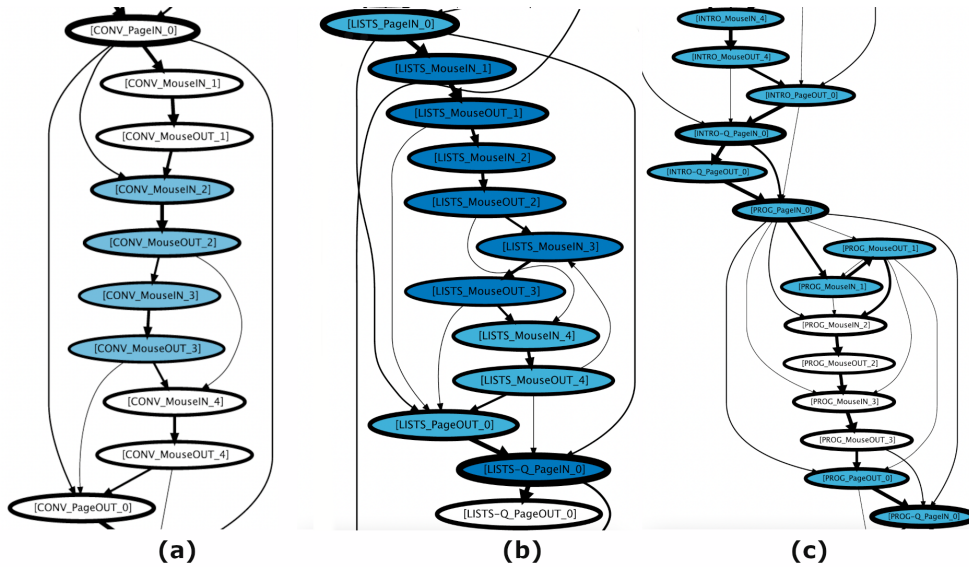


Fig. 6 The automated processes comparator analysis output of negative or positive outcomes, with respect to (a) trace frequency, (b) elapsed time, and (c) remaining time. Image generated with ProM. Full size image available at <https://bit.ly/3Tm6GJD>

4.3 Analysis of learning tracks

The three learning tracks

To explore the three learning paths, we consider the main measures on time and performance. As summarized in Table 5, *Track1* is longer than the other two (median duration of 42.1 minutes), achieving better results (71.2% of correct answers). On the opposite, *Track3* is relatively shorter (29.9 minutes), with a lower performance (64.9% of correct answer). These results suggest the existence of some differences, to be examined in more detail in the next paragraphs by focusing on time and outcome.

Table 5 Students' performances in the three tutorial tracks, i.e. the number of cases, the mean, the median, and the STD in terms of minutes for each track

Track	Cases	Median (min)	STD (min)	Correct answers
1	33%	42.1	22.4	71.2%
2	28%	32.4	26.6	70.1%
3	39%	29.9	35.2	64.9%

Time analysis of learning tracks

A further insight concerns the analysis of times between individual pages. We focus on the central activities of the tutorial concerning the three topics (of two lessons each) into which the flow described in Figure 1 is divided. As depicted by Figure 7, we

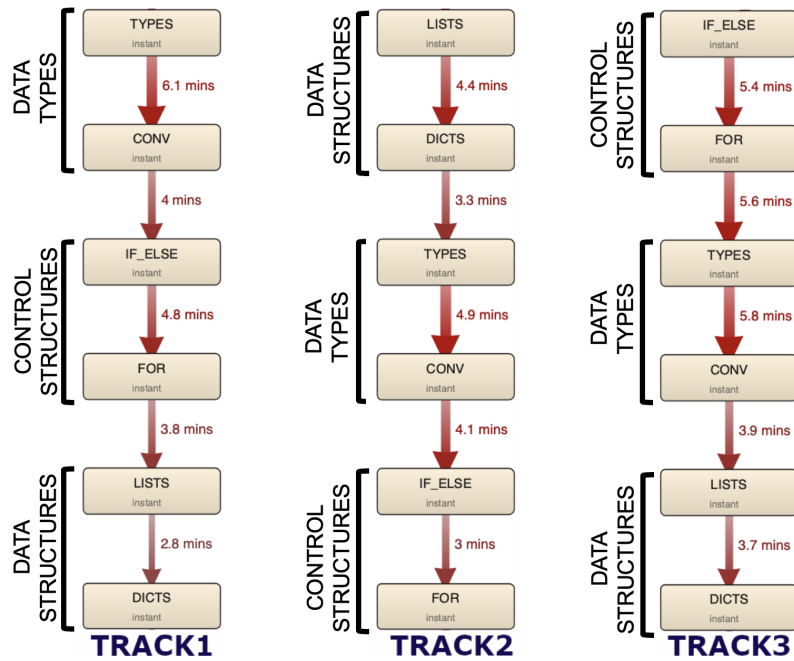


Fig. 7 Performance analysis (median duration between the activities) of the three tracks' central activities (the first part common to all tracks is not present). Image generated with Fluxicon DISCO. Full size image available at <https://bit.ly/3Tm6GJD>

examine the time between pages of the three tracks, i.e. the pairs TYPES and CONV (DATA TYPES topic), IF_ELSE and FOR (CONTROL STRUCTURES topic), LISTS and DICTS (DATA STRUCTURES topic).

Two interesting regularities appear relatively evident. First, we notice the regularity of a quickening towards the concluding activities in all tracks, regardless of track type. In fact, in each track, the initial topic always took longer than the others that follow in the exercise. Similarly, when the topic appears at the end of the track, it is always carried out faster. This phenomenon can be interpreted as a familiarity gained with the content of the tutorial or an indicator that the student gets bored and tries to go faster in the second part, regardless of the lessons he or she has to go through.

A second observation is that the order in which topics are presented affects the duration of the execution. Specifically, for the same content, the duration is different depending on whether it is presented earlier or later. For example, CONTROL STRUCTURES topics are performed more slowly if presented at the beginning (median duration of 5.4 minutes, in *Track3*) and much faster if presented later (2.8 minutes in *Track1* and 3 in *Track2*).

These recurrent activity flows, therefore, suggest presenting attention to the order of the activities, as the most important ones should be offered at the beginning of the short tutorial when attention appears highest.

Outcome analysis of learning tracks

A joint examination of the three tracks' median duration and the outcome provides additional insights. Positive cases are always longer than negatives for each track, as mentioned. More interestingly, the median duration of *Track1* is always higher than *Track2* or *Track3*, both for cases with positive (44.2 instead of 39 or 43.2) and negative outcomes (39.6 instead of 31.6 or 26.8). This seems to imply that *Track1* favors a greater depth of contents.

Focusing on *Track3*, students with positive outcome had a very long average duration, almost equal to *Track1*, while in contrast those with negative outcome were the group that went the fastest of all. A possible interpretation is that *Track3* forced those who wanted to achieve good results to pay more attention, while it accelerated the progress of the tutorial for those who were not motivated to achieve a good result.

The analysis of backward jumps (Table 6) confirms how *Track3* was the one that forced students to go deeper into the topics, regardless of whether the learning outcome is positive or negative. While the numerosity of the subgroups does not allow for generalization, these findings deserve to be further investigated, as they show the importance of focusing on the paths taken. A qualitative investigation would be necessary to understand the differences between the paths and evaluate the contents proposed by the learning track, which is out of the scope of the current work.

Table 6 Average number of jumps per page based on quiz result for the group of cases with positive outcome (OUT-POS) and negative outcome (OUT-NEG)

Cases	Track 1	Track 2	Track 3
OUT-POS	1.91	3.21	2.03
OUT-NEG	1.82	2.12	2.18

4.4 Outcome predictions

This Section describes the predictive models results to investigate the outcome of the pathway after the first part of the tutorial, according to our RQ2. As mentioned in Section 3.4, prefixes of lengths 40, 80, and 160 have been extracted to investigate the first half of the process whose outcome we want to predict. Table 7 describes the results obtained from the XGB and RF models. The prediction results improve as the prefix size increases. Apart from the shortest prefix length (40), which gets poor results, already with a length of 80, the XGB model (better than RF) gets results of some interest. Interestingly, XGB with IE is somehow always better than RF. In the best case, before the midpoint of the student's online course, the final trajectory was predicted with about 70% accuracy using XGB with IE coding (F1-Score of 0.6721, Accuracy of 0.6741, Precision of 0.6846, Recall 0.6781, and AUC 0.7221); both AUC and F1 are consistent in defining the best classifier for each prefix.

Even though the algorithms are both ensemble types (Dietterich (2000)), it can be observed that RF performs better with FE encoding while XGB with IE encoding. These prediction results are not only quite satisfactory in themselves, but more

importantly, they show a good possibility in our proof-of-concept, a sign that such an analysis can be done and at the same time provides a baseline from which to start and to compare with. In terms of time, the computation for training the machine learning models took about 30 minutes for the 40 prefixes to about 4.5 hours for the 160 prefixes (the most time-consuming optimization is that of the XGB).

Domain experts can analyse the prediction model's results and make other considerations. By using a prediction model, teachers can more timely identify students who might encounter difficulties in the tutorial. This allows them to intervene early and provide targeted support to improve students' performance. Knowing at-risk students allows teachers to adapt their teaching to meet the specific needs of these students. They can provide additional resources, offer individual tutoring sessions or change the pace of the course to ensure that at-risk students have a better chance of success. Focusing instructional resources on students needing additional support can optimise teaching efficiency. Teachers can allocate more time and resources to these students, enabling them to maximise their educational impact. Using the model as a continuous assessment tool, teachers can continuously monitor student performance throughout the tutorial. They can timely identify changes in students' performance over time and adapt teaching strategies accordingly. Moreover, by analysing the predictive data provided by the model, teachers can assess the effectiveness of their tutorial and identify areas requiring improvement. They can then modify course content, teaching methods, or assessments to maximise students' success.

5 Discussions

In this section, we discuss the strengths and weaknesses of our approach, some reflections on the capability of process mining, and the generalizability of our work.

Strengths and limitations of our work

The teachers involved in administering the tutorial evaluated the results positively. From an instructional point of view, information about the learning pathways allows teachers to understand what is happening within the specific lessons. Sequential learning has been recognized as the winning strategy in most cases. In addition, speed of execution and a lack of desire to go deeper were recognized as key factors in learning failure. We are aware that our study has some weaknesses. First, there is a lack of contextual knowledge, e.g., previous knowledge of programming skills from students involved in the tutorial (being non-computer courses, we assumed that almost everyone was ignorant on the topic - in any case, we are interested in an aggregate/average measure, so outliers are smoothed). Second, we do not discriminate between users with difficulties in using computers or interacting with technology. While we assumed they are a minority in our study, we will try to take this into account for future work. Third, our approach is focused on data that can be tracked by the information system. This means that the investigation of the cognitive dimension is not immediate. The qualitative analysis of the learner's educational context at the moment of the tutorial's administration is a common problem in other studies in the educational process

Table 7 Prediction results: for each prefix and its relative encoding (BE, IE or FE), it is possible to compute the performance (F1-score, Accuracy, Precision, Recall) of each algorithm (RF or XGB). The best performance is achieved by XGB in the configuration with prefixes encoded in IE mode, while the best RF results are obtained from prefixes encoded in FE mode. **Bold** values are the cases with the best results for each prefix length, while the second best ones are underlined

Algorithm	Prefix len	F1-score	Accuracy	Precision	Recall	AUC
RF (IE)	40	0.5677	0.5683	0.5702	0.5696	0.5788
RF (BE)	40	0.5522	0.5547	0.5586	0.5569	0.5777
RF (FE)	40	0.4223	0.4737	0.47	0.4816	0.5409
<u>XGB (IE)</u>	40	<u>0.5633</u>	<u>0.5652</u>	<u>0.5688</u>	<u>0.5672</u>	<u>0.594</u>
XGB (BE)	40	0.5305	0.5318	0.5341	0.5335	0.5465
XGB (FE)	40	0.559	0.5596	0.5615	0.5609	0.5878
RF (IE)	80	0.549	0.5823	0.6458	0.5956	0.6695
RF (BE)	80	0.546	0.5496	0.5574	0.5547	0.5886
RF (FE)	80	0.5598	0.5643	0.5743	0.5699	0.6443
XGB (IE)	80	0.6168	0.6184	0.626	0.6225	0.668
XGB (BE)	80	0.5534	0.5538	0.5565	0.5561	0.5515
<u>XGB (FE)</u>	80	<u>0.5895</u>	<u>0.5895</u>	<u>0.5907</u>	<u>0.5907</u>	<u>0.627</u>
RF (IE)	160	0.5479	0.5509	0.5571	0.5549	0.5875
RF (BE)	160	0.5699	0.5727	0.5798	0.5767	0.6026
RF (FE)	160	0.6182	0.6191	0.624	0.6219	0.6619
XGB (IE)	160	0.6721	0.6741	0.6846	0.6781	0.7221
XGB (FE)	160	0.49	0.4928	0.4964	0.4965	0.5267
XGB (FE)	160	0.5041	0.5225	0.5376	0.5308	0.5585

mining field. Finally, we can improve the survey by increasing data requested by the students, e.g. demographic data.

Finally, we offer some concluding remarks on the technology’s capacity adopted in this work. As the variation of events is relatively low, this has resulted in a limitation to the full utilisation of the process mining’s potential. Due to the lack of a wide variation of events, the insights generated in this work may not fully reflect the dynamics of the underlying processes. We have already pointed out that our work focuses on control flow analysis and the automatic extraction of events recorded in the computer system. This analysis may result in a narrow view of the process, potentially leading to incomplete or distorted conclusions, and must be incorporated with contextual knowledge, as mentioned above within the limitations of our study. To address this issue, we identify three main strategies that should be considered by future work aimed at leveraging PM techniques for this kind of analysis. First, there is a need for a diverse and comprehensive dataset. Future studies should aim to include a wider range of event types and instances to capture a broader spectrum of process variations; second, complement other analytical methods to provide a more holistic understanding of the process, e.g. through qualitative analysis; third, case selection should be carefully considered, including a diverse sample of cases in order to improve the applicability of PM and lead to more robust results; fourth, an iterative and integrated approach with domain experts (as suggested by studies on interactive process mining) starting from

the preliminary data collection and analysis stages to gradually improve the richness of the dataset.

Generalizability of the results

Regarding the approach’s generalizability, we highlight that the methodology proposed to generate the event log can be easily applied to leverage process mining on other web-based tutorials under the condition that they track similar kinds of data. We argue that such conditions are easy to satisfy. Our work is based on web technologies, which became a common way to offer self-learning tutorials. In addition, our results, intended as data, techniques and instruments are publicly accessible, thus the results can be replicated. Second, the proposed solution can be easily applied to a broader context, both as a type of user of the tutorial and as content. Short-term tutorials, in fact, can be adopted for various types of audiences, not only university students, as in our case. Furthermore, the contents can also vary, defining in a congruent way an adequate linguistic register for the description of the proposed contents.

6 Related work

This section provides an overview of related work, highlighting the main differences in our work to position it with respect to the state of the art.

Our work falls within the stream of studies on learning with computer-based methods, which typically involve the measurement, collection, analysis, and reporting of data about learners and the context in which they occurs. Such studies investigate students’ actions through traces detected by e-learning systems in the context of LMSs (Turnbull et al (2020)). Courses based on Learning Management Systems (LMSs), such as Massive Open Online Courses (MOOCs). In the following, we focus on work leveraging process mining and machine learning techniques to model learning processes and predict their outcomes.

Learning processes and process mining. The recent discipline of PM concerns ideas, methods, and tools to extract knowledge from a time series of activities, i.e. event logs (van der Aalst (2016)). The students’ behavior can be explored in three directions: comparison of students’ behavior, performance prediction based on students’ behavior, and learning strategy evaluation (Wafda et al (2022)). Several previous studies already explored PM to improve educational processes (Ghazal et al (2017)). Process discovery techniques were used also to investigate students’ different web behavior strategies in tackling quizzes in online tests (Juhaňák et al (2019)). Similar to our work, the authors investigated the adoption of PM to analyse students’ quiz-taking behavior patterns, but they focused on an LMS. In Moreno et al (2021), the authors promoted a correlation study between the behavior of the learner (i.e. the number of connections between the sections of a course followed) during the learning process and their mark obtained on the final exam, starting from an event log obtained from the LMS. The study in Cerezo et al (2020) aims to discover the self-regulated learning processes of students in an e-learning course using PM techniques, by applying the Inductive Miner algorithm to interaction traces from 101 university students on the Moodle platform. The algorithm revealed optimal models for both passing and failing students, offering

insights into successful self-regulated learning processes. Another study used process mining from an university LMS to analyse learners' behavior (Sedrakyan et al (2016)), while the study in (Sedrakyan et al (2014)) analyses 20 cases to study patterns linked to learning performance, enhancing teaching guidance with process-oriented feedback.

Predicting the learning outcome. In (Yu and Jo (2014)), web-log data from a Moodle-based LMS were used to investigate 84 students' academic achievements. A multi-regression analysis showed a significant correlation with the final learning grade. Finally, the authors suggest that "educators should pay more attention to improve the process of learners' achievement". In another study, students' behavior has been monitored for evaluation purposes during a semester by constructing an event log of their activities in a specific LMS (Cenka et al (2022)). The authors stated that teachers must design teaching strategies that provide early or real-time detection of students who do not follow the learning path. Predictive models have been implemented using students' behavior based on an edX-based LMS (Deeva et al (2022)), to identify underperforming students early (De Smedt et al (2019)), as well as students' abilities before and after problem-solving tasks (Liu et al (2022)) by using Gradient Boosting Decision Trees on historical event logs. Other predictive studies involve the automated analysis of traces left by students in MOOCs (Romero and Ventura (2020)), also by differentiating various subgroups of learners (Luna et al (2022)), demonstrating how to predict the performance of students at an early stage (Umer et al (2017)), as well as to predict student's outcome in a course by exploiting information on LMSs (Umer et al (2019)). A previous research identified three main types of outcome prediction: the exact final grade (e.g., the range can correspond to a scale from 0 to 10), a mapping into a limited number of categories, usually 4 or 5, or a discretization into two categories, i.e. negative/positive (Hu et al (2017)). Our work focused on the last categorization.

Learning styles. Learning styles have been the subject of many studies that recognized the existence of multiple factors, often attributable to the learner's personal characteristics or the used technologies. A recent literature review summarized the existing theories on learning styles (Truong (2016)). As they generally suffer from validity and reliability issues (Coffield et al (2004)), no theory outweighs the others. Nevertheless, one of the most popular theories that has been applied in e-learning systems is the Felder-Silverman one (Felder and Silverman (1988)). Their theory includes the categorization between sequential versus global learning styles: sequential learning style concerns the acquisition of understanding in a linear fashion, with a logical progression of ordered steps; on the contrary, a global learning style involves absorbing material more disorderedly, including non-linear connections and jumps between the various parts (Felder and Brent (2016)).

In Mukala et al (2015), process discovery has been applied to investigate learning styles in a MOOC course, finding a positive correlation between sequential learning and students' performance. Process analysis revealed that successful students followed the learning path while less successful students did not (Cenka et al (2022)). A relevant issue concerns the consideration of the learners' goals and their regulatory mechanisms. A conceptual model and a practical case example have been proposed with the adoption of a feedback-driven dashboard, i.e. a dashboard designed on the basis of empirical evidence to enhance learning regulation by providing both cognitive and behavioral

feedback (Sedrakyan et al (2014)). In their work, process discovery has been adopted to investigate the interactions between user participants. Previously, process discovery has been used to analyze the detailed logs of novice users' interactions within a specific tool in Sedrakyan et al (2020). This kind of research connects process-mining enabled analysis of learning processes and behaviors with learning theories, aligning data collection and analysis with underlying learning processes from the learning sciences. By examining 20 cases with over 10,000 logged events, process discovery helped identify patterns and sequences in the learning process. Our work contributes to this cross-domain direction by studying learning behavior in a real-world situation.

The study in Liegle and Janicki (2006) explores how customizing web-based learning to match individual styles -distinguishing between "Explorers" (who prefer self-navigation) and "Observers" (who follow structured paths)- can enhance learning effectiveness. With 58 participants, findings suggest that learning outcomes improve when the system's navigation style aligns with the user's learning preference, emphasizing the potential of adaptive learning platforms. "Explorers" performed better when jumping between content, while "Observers" excelled with linear navigation. This indicates that customized learning platforms, responsive to individual preferences, can enhance learning outcomes. We investigated these modes of behavior in a short tutorial. Finally, a relevant feature of a learning style concerns its duration. Our assumption is that the learning style remains fixed for the duration of the tutorial, according to Truong (2016).

Learning design. Studies on learning styles demonstrated how hypermedia technologies benefit learners with different needs (Liu and Reed (1994)). As in our work, the application of automated process analysis in education has also been shown to have impact on the field of Learning Design, which can be defined as "a methodology for enabling teachers/designers to make more informed decisions in how they go about designing learning activities and interventions, which is pedagogically informed and makes effective use of appropriate resources and technologies" (Macfadyen et al (2020)).

According to a recent review, the most frequent kind of learning concerns 'assimilative activity', such as reading module materials, which corresponds to the one addressed in our work (Rienties et al (2015)).

Our study leverages process mining and machine learning techniques with a similar purpose than previous studies, namely, to determine learning processes describing behaviors of successful and less successful students and to predict students' performance before the end of the learning trajectory. Compared with the state-of-the-art, the distinguishing features and improvements of our work include the following main points:

- the focus of our analysis concerns a learning path of short duration (two hours at most) and not months or years as in most studies;
- the exploitation of web technologies to track behavior within tutorial paragraphs on web pages, and do not use data from pre-existing systems such as MOOCs used by most studies in this area;
- the application of process mining analysis on short tutorials in such a tracking system.

To the best of our knowledge, no previous work has addressed this type of analysis on relatively short learning paths, exploiting web-based technologies with process mining techniques.

7 Conclusions

The paper proposed a methodology for studying the learning of short tutorials using the combination of a web tracking system and the application of process mining techniques at descriptive level. In a practical case study, we demonstrated how this methodology could investigate the learning path and activity flows of students who did well and poorly (RQ1). Our analysis suggests differences in students' learning and satisfaction adopting a specific order among topics.

Finally, the proposed methodology can be applied to identify possible bottlenecks and other hints in relatively short learning paths. The fact that the student who performs poorly goes fast from the start, as well as behaves with a more linear path instead of jumping back to previous paragraphs, may suggest that the system can make appropriate slowdowns or alerts when it detects potentially dangerous behavior in learning. The prediction results (RQ2) encourage the adoption of a prediction system in the tutorial's initial part (ideally at the end of the third lesson) to investigate students who are at risk of insufficient learning after the first part of the course.

Future work. We aim to increase the number of tutorial administrations to obtain more statistically significant results. In addition, we plan to extend the survey with more variables, e.g., demographic data and previous knowledge. From a learning design perspective, we would like to gather more suggestions on the usability front and address bottleneck analysis of the present tutorial to identify valuable suggestions for implementing an improved version. The new version of the tutorial can then be resubmitted to another similar set of students to investigate the improvements, as part of prescriptive process monitoring (Kubrak et al (2022)). For instance, our PM analysis can identify paragraphs of the actual version of the tutorial where most students spend too much and be grounds for restructuring for a new, improved version. We aim to extend our work by implementing appropriate feedback to students, in order to investigate aspects of the cognitive thinking process or regulation.

Moreover, as domain experts have suggested, we may include a survey of student's initial knowledge of the subject (programming in Python) before the tutorial to assess its benefits at the end of the learning path. As far as the prediction phase is concerned, in future research, we intend to explore explainability issues (Meo et al (2022)) as well as deep-learning models such as long short-term memory, generative adversarial networks, and transformers, which require larger amounts of data to be trained effectively (Jordan and Mitchell (2015)).

Declarations

Ethical Approval

Not Applicable

Availability of supporting data

The datasets and scripts generated for this research are available in the Google Drive repository, <https://bit.ly/3Tm6GJD>

Competing interests

I declare that the authors have no competing interests as defined by Springer, or other interests that might be perceived to influence the results and/or discussion reported in this paper.

Funding

Not Applicable

Authors' contributions

R.N. and E.S. wrote the main manuscript; L.G. carried out the supervision of the PM results.

Acknowledgements

The research work in this article was partially conducted as part of the Circular Health European Digital Innovation Hub (CHEDIH) project.

References

- van der Aalst WMP (2016) Process Mining - Data Science in Action, Second Edition. Springer, <https://doi.org/10.1007/978-3-662-49851-4>
- Bergstra J, Komer B, Eliasmith C, et al (2015) Hyperopt: a python library for model selection and hyperparameter optimization. Computational Science & Discovery 8(1):014008. <https://doi.org/10.1088/1749-4699/8/1/014008>
- Bogarín A, Cerezo R, Romero C (2018) A survey on educational process mining. WIREs Data Mining and Knowledge Discovery 8(1). <https://doi.org/10.1002/WIDM.1230>
- Bolt A, de Leoni M, van der Aalst WMP (2018) Process variant comparison: Using event logs to detect differences in behavior and business rules. Information Systems 74(Part):53–66. <https://doi.org/10.1016/J.IS.2017.12.006>
- Buckland MK, Gey FC (1994) The relationship between recall and precision. Journal of the American Society for Information Science 45(1):12–19. [https://doi.org/10.1002/\(SICI\)1097-4571\(199401\)45:1](https://doi.org/10.1002/(SICI)1097-4571(199401)45:1)
- Bujlow T, Carela-Español V, Solé-Pareta J, et al (2017) A survey on web tracking: Mechanisms, implications, and defenses. Proceedings of the IEEE 105(8):1476–1510. <https://doi.org/10.1109/JPROC.2016.2637878>

- Cenka BAN, Santoso HB, Junus K (2022) Analysing student behaviour in a learning management system using a process mining approach. *Knowledge Management & E-Learning* 14(1):62–80. <https://doi.org/10.34105/j.kmel.2022.14.005>
- Cerezo R, Bogarín A, Esteban M, et al (2020) Process mining for self-regulated learning assessment in e-learning. *Journal of Computing in Higher Education* 32(1):74–88. <https://doi.org/https://doi.org/10.1007/s12528-019-09225-y>
- Coffield F, Ecclestone K, Hall E, et al (2004) Learning styles and pedagogy in post-16 learning: A systematic and critical review. Learning and Skills Research Council, London
- De Smedt J, Deeva G, De Weerd J (2019) Mining behavioral sequence constraints for classification. *IEEE Transactions on Knowledge and Data Engineering* 32(6):1130–1142. <https://doi.org/10.1109/TKDE.2019.2897311>
- Deeva G, Smedt JD, Saint-Pierre C, et al (2022) Predicting student performance using sequence classification with time-based windows. *Expert Systems with Applications* 209:118182. <https://doi.org/https://doi.org/10.1016/j.eswa.2022.118182>
- Di Francescomarino C, Ghidini C (2022) Predictive process monitoring. In: van der Aalst WMP, Carmona J (eds) *Process Mining Handbook, Lecture Notes in Business Information Processing*, vol 448. Springer, p 320–346, https://doi.org/10.1007/978-3-031-08848-3_10
- Dietterich TG (2000) Ensemble methods in machine learning. In: Kittler J, Roli F (eds) *Multiple Classifier Systems, First International Workshop, MCS 2000, Cagliari, Italy, June 21-23, 2000, Proceedings, Lecture Notes in Computer Science*, vol 1857. Springer, pp 1–15, https://doi.org/10.1007/3-540-45014-9_1
- Ermakova T, Fabian B, Bender B, et al (2018) Web tracking - A literature review on the state of research. In: Bui T (ed) *51st Hawaii International Conference on System Sciences, HICSS 2018, Hilton Waikoloa Village, Hawaii, USA, January 3-6, 2018. ScholarSpace / AIS Electronic Library (AISeL)*, pp 1–10, URL <https://hdl.handle.net/10125/50485>
- Fawcett T (2006) An introduction to roc analysis. *Pattern Recognition Letters* 27(8):861–874. <https://doi.org/https://doi.org/10.1016/j.patrec.2005.10.010>
- Felder R, Silverman L (1988) Learning and teaching styles in engineering education. *Journal of Engineering Education* 78:674–681
- Felder RM, Brent R (2016) *Teaching and learning STEM: A practical guide*. John Wiley & Sons
- Feng G, Fan M (2024) Research on learning behavior patterns from the perspective of educational data mining: Evaluation, prediction and visualization. *Expert Systems*

- with Applications 237:121555. <https://doi.org/https://doi.org/10.1016/j.eswa.2023.121555>
- Géron A (2022) Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow. O'Reilly Media, Inc.
- Ghazal MA, Ibrahim O, Salama MA (2017) Educational process mining: A systematic literature review. In: 2017 European Conference on Electrical Engineering and Computer Science (EECS), pp 198–203, <https://doi.org/10.1109/EECS.2017.45>
- Hu X, Cheong CWL, Ding W, et al (2017) A systematic review of studies on predicting student learning outcomes using learning analytics. In: Hatala M, Wise AF, Winne P, et al (eds) Proceedings of the Seventh International Learning Analytics & Knowledge Conference, Vancouver, BC, Canada, March 13-17, 2017. ACM, pp 528–529, <https://doi.org/10.1145/3027385.3029438>
- Jordan MI, Mitchell TM (2015) Machine learning: Trends, perspectives, and prospects. *Science* 349(6245):255–260. <https://doi.org/10.1126/science.aaa8415>
- Juhaňák L, Zounek J, Rohlíková L (2019) Using process mining to analyze students' quiz-taking behavior patterns in a learning management system. *Computers in Human Behavior* 92:496–506. <https://doi.org/https://doi.org/10.1016/j.chb.2017.12.015>
- Kabudi T, Pappas IO, Olsen DH (2021) Ai-enabled adaptive learning systems: A systematic mapping of the literature. *Computers and Education: Artificial Intelligence* 2:100017. <https://doi.org/10.1016/J.CAEAI.2021.100017>
- Kubrak K, Milani F, Nolte A, et al (2022) Prescriptive process monitoring: *Quo vadis?* *PeerJ Comput Sci* 8:e1097. <https://doi.org/10.7717/PEERJ-CS.1097>
- Liegle JO, Janicki TN (2006) The effect of learning styles on the navigation needs of web-based learners. *Computers in Human Behavior* 22(5):885–898. <https://doi.org/https://doi.org/10.1016/j.chb.2004.03.024>
- Liu F, Zhao L, Zhao J, et al (2022) Educational process mining for discovering students' problem-solving ability in computer programming education. *IEEE Transactions on Learning Technologies* 15(6):709–719. <https://doi.org/10.1109/TLT.2022.3216276>
- Liu M, Reed W (1994) The relationship between the learning strategies and learning styles in a hypermedia environment. *Computers in Human Behavior* 10(4):419–434. [https://doi.org/https://doi.org/10.1016/0747-5632\(94\)90038-8](https://doi.org/https://doi.org/10.1016/0747-5632(94)90038-8)
- Luna JM, Fardoun HM, Padillo F, et al (2022) Subgroup discovery in moocs: a big data application for describing different types of learners. *Interactive Learning Environments* 30(1):127–145. <https://doi.org/10.1080/10494820.2019.1643742>

- Macfadyen LP, Lockyer L, Rienties B (2020) Learning design and learning analytics: Snapshot 2020. *Journal of Learning Analytics* 7(3):6–12. <https://doi.org/https://doi.org/10.18608/jla.2020.73.2>
- Maggi FM, Francescomarino CD, Dumas M, et al (2014) Predictive monitoring of business processes. In: Jarke M, Mylopoulos J, Quix C, et al (eds) *Advanced Information Systems Engineering - 26th International Conference, CAiSE 2014, Thessaloniki, Greece, June 16-20, 2014. Proceedings, Lecture Notes in Computer Science*, vol 8484. Springer, pp 457–472, https://doi.org/10.1007/978-3-319-07881-6_31
- Meo R, Nai R, Sulis E (2022) Explainable, interpretable, trustworthy, responsible, ethical, fair, verifiable AI... what's next? In: Chiusano S, Cerquitelli T, Wrembel R (eds) *Advances in Databases and Information Systems - 26th European Conference, ADBIS 2022, Turin, Italy, September 5-8, 2022, Proceedings, Lecture Notes in Computer Science*, vol 13389. Springer, pp 25–34, https://doi.org/10.1007/978-3-031-15740-0_3
- Moreno M, Exposito E, Gueye M (2021) Process mining model to visualize and analyze the learning process. In: REES AAEE 2021 conference: Engineering Education Research Capability Development: Engineering Education Research Capability Development, Engineers Australia Perth, WA, pp 698–706, <https://doi.org/10.52202/066488-0083>
- Mukala P, Buijs JCAM, Leemans M, et al (2015) Learning analytics on coursera event data: A process mining approach. In: Ceravolo P, Rinderle-Ma S (eds) *Proceedings of the 5th International Symposium on Data-driven Process Discovery and Analysis (SIMPDA 2015)*, Vienna, Austria, December 9-11, 2015, *CEUR Workshop Proceedings*, vol 1527. CEUR-WS.org, pp 18–32, URL <https://ceur-ws.org/Vol-1527/paper2.pdf>
- Nai R, Sulis E, Marengo E, et al (2023) Process mining on students' web learning traces: A case study with an ethnographic analysis. In: Viberg O, Jivet I, Muñoz-Merino PJ, et al (eds) *Responsive and Sustainable Educational Futures - 18th European Conference on Technology Enhanced Learning, EC-TEL 2023, Aveiro, Portugal, September 4-8, 2023, Proceedings, Lecture Notes in Computer Science*, vol 14200. Springer, pp 599–604, https://doi.org/10.1007/978-3-031-42682-7_48
- Rienties B, Toetenel L, Bryan A (2015) "scaling up" learning design: impact of learning design activities on LMS behavior and performance. In: Baron J, Lynch G, Maziarz N, et al (eds) *Proceedings of the Fifth International Conference on Learning Analytics And Knowledge, LAK '15, Poughkeepsie, NY, USA, March 16-20, 2015*. ACM, pp 315–319, <https://doi.org/10.1145/2723576.2723600>
- Rizzi W, Francescomarino CD, Ghidini C, et al (2022) Nirdizati: an advanced predictive process monitoring toolkit. *CoRR* abs/2210.09688. <https://doi.org/10.48550/ARXIV.2210.09688>

- Romero C, Ventura S (2020) Educational data mining and learning analytics: An updated survey. *WIREs Data Mining Knowl Discov* 10(3). <https://doi.org/10.1002/widm.1355>
- Sedrakyan G, Snoeck M, De Weerd J (2014) Process mining analysis of conceptual modeling behavior of novices—empirical study using jmermaid modeling and experimental logging environment. *Computers in Human Behavior* 41:486–503. <https://doi.org/https://doi.org/10.1016/j.chb.2014.09.054>
- Sedrakyan G, De Weerd J, Snoeck M (2016) Process mining enabled feedback: ”tell me what i did wrong” vs. ”tell me how to do it right”. *Computers in Human Behavior* 57:352–376. <https://doi.org/https://doi.org/10.1016/j.chb.2015.12.040>
- Sedrakyan G, Malmberg J, Verbert K, et al (2020) Linking learning behavior analytics and learning science concepts: Designing a learning analytics dashboard for feedback to support learning regulation. *Computers in Human Behavior* 107:105512. <https://doi.org/https://doi.org/10.1016/j.chb.2018.05.004>
- Sweller J (1994) Cognitive load theory, learning difficulty, and instructional design. *Learning and Instruction* 4(4):295–312. [https://doi.org/https://doi.org/10.1016/0959-4752\(94\)90003-5](https://doi.org/https://doi.org/10.1016/0959-4752(94)90003-5)
- Truong HM (2016) Integrating learning styles and adaptive e-learning system: Current developments, problems and opportunities. *Computers in Human Behavior* 55:1185–1193. <https://doi.org/10.1016/j.chb.2015.02.014>
- Turnbull D, Chugh R, Luck J (2020) Learning management systems: An overview. In: Tatnall A (ed) *Encyclopedia of Education and Information Technologies*. Springer International Publishing, p 1–7, https://doi.org/10.1007/978-3-319-60013-0_248-1
- Umer R, Susnjak T, Mathrani A, et al (2017) On predicting academic performance with process mining in learning analytics. *Journal of Research in Innovative Teaching & Learning* 10(2):160–176. <https://doi.org/https://doi.org/10.1108/JRIT-09-2017-0022>
- Umer R, Mathrani A, Susnjak T, et al (2019) Mining activity log data to predict student’s outcome in a course. In: *Proceedings of the 2019 International Conference on Big Data and Education*. Association for Computing Machinery, New York, NY, USA, ICBDE ’19, p 52–58, <https://doi.org/10.1145/3322134.3322140>
- von Rosing M, White S, Cummins F, et al (2015) Business process model and notation—BPMN. <https://doi.org/https://doi.org/10.1016/B978-0-12-799959-3.00021-5>
- Wafda F, Usagawa T, Mahendrawathi E (2022) Systematic literature review on process mining in learning management system. In: *2022 IEEE International Conference on Industry 4.0, Artificial Intelligence, and Communications Technology (IAICT)*, pp

160–166, <https://doi.org/10.1109/IAICT55358.2022.9887428>

Yu T, Jo I (2014) Educational technology approach toward learning analytics: relationship between student online behavior and learning performance in higher education. In: Pistilli MD, Willis J, Koch D, et al (eds) Learning Analytics and Knowledge Conference 2014, LAK '14, Indianapolis, IN, USA, March 24-28, 2014. ACM, pp 269–270, <https://doi.org/10.1145/2567574.2567594>