Understanding Student Behavior Using Active Window Tracking and Process Mining Contribution Title

Mahendrawathi ER[1][0000-0002-5966-2274] Wouter van der Waal[2][0000-0001-6514-7570]
Iris Beerepoot[2][0000-0002-6301-9329] M. Aqmal R.R Putra[1][0009-0009-5740-5686]
Hardhika Propitadewa[1][0009-0006-1751-0877]

[1] Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia
mahendra_w@is.its.ac.id, aqmal2009@gmail.com,
propitadewa@gmail.com,
[2]Utrecht University, The Netherlands
i.m.beerepoot@uu.nl,w.g.vanderwaal@uu.nl

Abstract. This paper proposes a new way of collecting and processing event logs using Active Window Tracking (AWT) to investigate media multitasking (MMT) among students in higher education institutions in Indonesia. Students recorded their computer windows while doing assignments and midterms. Data from the students were preprocessed and structured into event logs. Correlation analysis indicated that MMT has no direct correlation with performance. The PM results revealed that students engaging in MMT frequently switch between assignments, social media, and multimedia. High-scoring students focused more on assignment-related activities, while low-scoring students started late, multitasked extensively, and submitted their work close to the deadline. While these results indicate that MMT does not directly affect the student's performance for the type of assignment, MMT extends work duration. Students tend to work closer to the deadline, so they often work very late into the night, negatively impacting their well-being. Recommendations are provided to mitigate these issues.

Keywords: media multitasking, student behavior, active window tracking, event log, process mining.

1 Introduction

The learning process has dramatically changed with the advancement of technology. Learning Management Systems (LMS) allow all students' learning activities and academic achievements to be documented in an integrated system, providing valuable data for lecturers [7]. Educational Data Mining (EDM) is a field aimed at extracting and analyzing this data. However, understanding the learning process requires a more comprehensive, process-centric approach called Education Process Mining (EPM). EPM uses algorithms to find patterns in event log data from educational systems, providing insights into actual learning activities [1, 6].

The ability of students to manage tasks and focus affects their academic performance and mental health [18, 23]. Students often face multiple tasks within limited time frames, making it hard to focus, especially with increased media access and multitasking [17, 21]. Media multitasking (MMT), used as a coping strategy, burdens working memory and cognitive function, reducing focus and academic performance [16, 17, 20]. The theory of attention highlights that individuals have limited cognitive resources, and multitasking increases cognitive load, hindering information retention and understanding [22]. It is suggested that focusing on one task can improve learning outcomes by reducing the working memory load and enhancing effective learning [11].

Recent studies have analyzed student behavior using process mining (PM) [9, 13–15]. However, no study has explored the use of media multitasking during the learning process, such as doing assignments or taking exams, which is an interesting area to explore. The investigation of MMT behavior requires a new approach to collecting and creating an event log. This study will use Active Window Tracking (AWT) [5] to gain insights into students' media multitasking behavior during assignments and midterms. AWT records user interactions to create detailed event logs , which will be analyzed with PM and correlation analysis to understand student behavior and its impact on academic performance.

The remainder of this paper is structured as follows. Relevant literature on media multitasking and EPM is discussed in section 2. Section 3 presents the method used in the paper to collect and process the event log from the AWT records. Results from the PM and correlation analysis are presented in section 4. A discussion on the implications of our study for educators in section 5 follows this. Concluding remarks are given in the final section.

2 Related Work

2.1 Educational Data Mining and Educational Process Mining

Educational Data Mining (EDM) develops methods to explore educational data, applying data mining techniques to address key questions in education [4, 20]. Studies show EDM's value in extracting insights into student behavior and learning effectiveness, including predicting performance, detecting unwanted behavior, and grouping students [3, 12, 19]. Educational Process Mining (EPM) is an emerging area within Educational Data Mining (EDM) that aims to reveal implicit knowledge and enhance understanding of educational processes. EPM uses log data from educational environments to uncover, analyze, and visually describe these processes [6]. Specifically, EPM applies Process Mining (PM) techniques to raw educational data [6]. Several studies have investigated student behavior in the learning process with PM. For example, PM has been used to analyze student behavior in completing quizzes on the LMS [13]. Student behavior in accessing the LMS has also been examined by [9], as well as the relationship between LMS log data and student achievement [14]. However, using LMS data alone does not allow for analyzing the full breadth of student behavior, e.g., other tools they are using while doing assignments. In this study, we will record this behavior using AWT, allowing for extended behavioral analysis, for example exploring multitasking behavior.

2.2 **Multitasking and Media Multitasking.**

Multitasking involves performing multiple tasks simultaneously through rapid attention switching, dividing attention, or task planning, which is common in daily activities like meeting deadlines or managing chores [8]. Media multitasking (MMT), a subset, involves using multiple digital media streams at once, such as texting while watching TV or checking social media while gaming. This behavior, especially prevalent among adolescents and young adults, has increased with smartphone use.

Despite its perceived benefits, MMT negatively affects cognitive performance. It increases cognitive load due to frequent attention switching, disrupts memory and learning, and impairs executive functions like planning and self-regulation. Long-term effects are still under study, but MMT is linked to mental health issues such as depression and anxiety and may impact cognitive development in children and adolescents [20]. The impact of MMT on homework and academic performance has been studied by [15]. To our knowledge, Active Window Tracking has not been used to study multimedia multitasking, and no study has utilized PM to observe MMT behavior during assignments or midterms.

3 Method

3.1 Context

Data is taken from the students' activities in two different batches that took a Business Process Modeling course at a public university in Surabaya, Indonesia. The students are in their fourth semester. Batch 1 consists of class 1, with 30 students, and class 2, with 31 students. Batch 2 has two classes with 17 students each. The BPM course is supported by a Learning Management System (LMS). The lecturer used the LMS to provide course materials, announce and collect assignments and conduct assessments.

3.2 Event Log Collection and Processing

Three learning activities, including two assignments and a midterm, are used as datasets from each batch. Students must create BPMN models based on textual descriptions for assignments 1 (A1) and 2 (A2). The assignments are delivered via the LMS, and students are given certain deadlines to submit them. The students can do the assignments outside the class hours. For batch 1, A1 is announced on 18/09/2023, 6.00 AM and the deadlines for the two classes are on 20/09/2023, 6.00 AM, while A2 is announced on 26/09/2023, 6.00 AM, and the deadlines are 2/10/2023, 6:00 AM. Only A2 in batch 2 has different deadlines. Details about the assignments and deadlines can be found in supplementary materials. After the submission deadline, the lecturers discussed the assignments during the class. The mid-term is conducted during in-person meetings to assess their modeling skills.

To collect the event log, students in the four classes are asked to record their windows using the Active Window Tracking application, tockler.io, while doing their assignments and mid-term. The experiments are conducted in line with the ethical procedure of Institut Teknologi Sepuluh Nopember. Students' involvement in the experiment is voluntary, and the students who volunteer fill in a formal consent form. Even after volunteering, students can opt-out (i.e., deciding not to submit their tockler.io recording) whenever they feel uncomfortable. Students who continue the experiments submit their Tockler log results in the LMS.

After all data was collected, we started event log processing. Python scripts were used for the data pre-processing stage, which involved tagging and labeling the window titles recorded on Tockler based on pre-defined keywords. For example, if the window title contains keywords related to Spotify (a music player application), the code will be “multimedia”. Labeling was done to make it easier for the researchers to identify the activities carried out by participants while completing the assignment or midterm. The complete list of labels and labeling process can be found in the supplementary materials.

Data cleaning is also done to eliminate double logging generated by preprocessing. To facilitate participant identification and analysis with PM, a new attribute called User ID is added to each participant’s log data file. Each User ID uniquely identifies the log data associated with each participant. After adding the User ID attribute, each participant's activity recording data is combined into one event log. The event log workflow will combine the three core attributes for each recorded activity: (1) *Case ID*: An attribute that acts as a *unique identifier* for each participant (in this case, the User ID), (2) *Activity*: represents a label assigned based on the content of the window title (e.g., "Social Media", "Modeling tools"), and (3) *Timestamp*: Records the specific time when each activity occurs. An example of a log event structure from results preprocessing activity recording log data for one participant (one Case ID) can be found in supplementary material.

3.3 Correlation Analysis and Process Mining.

To help identify interesting patterns for our process mining analysis, we first conducted a correlation analysis on several variables: the number of activities, the duration students take to do the assignment or midterm, and the score/grade. The activities are classified into those related to assignments (related) and those unrelated (unrelated) to assignments, such as doing other assignments, accessing social media, zoom, or other tasks. An additional variable called MMT density is calculated by dividing the number of activities by the duration. MMT measures the intensity of the windows switching within a certain period. Additionally, we analyze two temporal variables to calculate the intervals when students submit their assignment with 1) the official submission deadline and 2) midnight preceding the deadline. These variables offer critical insights into student behavior, potentially highlighting trends in time management and last-minute work habits that eventually correspond to their well-being. For the correlation analysis, we used Spearman correlations, as most individual variables were not normally distributed. We used a Bonferroni-Holm correction to account for the high number of comparisons. PM with the Apromore tool is then used to investigate interesting patterns based on correlation analysis.

4 Results

4.1 Media Multitasking and its Impact

The correlation analysis result in **Fig. 1** reports only the statistically significant correlations, all with a rounded p-value of 0.000. First, we look at the student's behavior related to activities and MMT density. For both assignments, **the duration is negatively correlated with the MMT, positively with the number of related activities**, and **strongly correlated with the number of unrelated activities**. Then, we also found correlations between the number of related activities to the number of unrelated activities within both assignments. Students who perform more relevant activities also tend to perform more irrelevant ones. Although the effect is weak, students who perform more relevant activities for A1 also tend to do that for A2. A1’s unrelated activities seem to be weakly correlated with the number of both types of activities in A2. These results indicate that students tend to have the same behavior while doing both assignments. Using PM, we further investigate the connection between these related and unrelated tasks in Section 4.2.

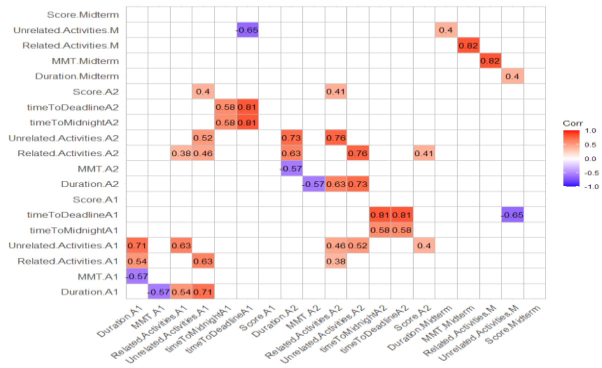

Fig. 1. Correlation results. Only the statistically significant correlation values are shown.

We also want to know whether students’ MMT affects their performance. With A2, we find that the more relevant activities a student performs, the higher their expected score. Interestingly, nothing indicates that MMT has a clear relation with scores. The relationship between MMT and performance for A1 is close to significant. For A2 and the midterm, we find nothing close to significance with low correlation. The strongest correlation is between the midterm's MMT and the number of related activities. As this observation does not align with the established theory about MMT, we investigate the relationship between MMT and performance with process mining in Section 4.2.

The time to midnight for A1 is weakly correlated with the hand-in times of A2. We find a much stronger effect if we check the same correlation for the time to the deadline of A1 and the hand-in times of A2. Finally, we found a moderately strong negative effect between the time to the deadline of A1 and the number of unrelated activities of the midterm. Students who handed in their work closer to the deadline seemed to perform more unrelated activities during the midterm.

4.2 Insights from Process Mining (PM)

Results from the correlation analysis help us identify how the variables relate to one another, which needs to be investigated further. We use PM to provide deeper insights into some of the strong correlations and unexpected findings to understand students' media multitasking.

Activities, Duration and Media Multitasking Intensity.

The correlation analysis result indicates that students who spend more time on an assignment perform more activities, especially those unrelated to the task. They switch less between screens, as the negative correlation with MMT shows. This could mean students who take more time are more focused or may have been idle or not actively involved with the assignment.

To investigate these correlations, we filter one case from Batch 1 with the longest duration during assignment one from Class 1. As shown in **Fig. 2**, this student activates Tockler for two days, with several idle times in between. This may be because, in the first assignment, the student was unfamiliar with AWT and did not change the setting. Tockler is immediately launched whenever the student activates their laptops. This way, the students continuously recorded their activities.

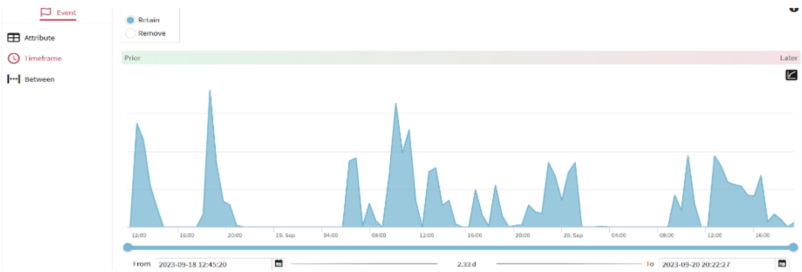

Fig. 2. Timeframe of Student from Batch 1 with the Longest Duration

To exclude those students who continuously recorded their activities, we focused on Batch 1, Class 1 while doing A2, as the students should be more familiar with operating AWT. We then drilled down on students who could finish in less than 2 hours, assuming that these students deliberately recorded their activities when they were about to do their assignments. There are 12 out of 30 students that belong to this category. The process map is shown in **Fig. 3**.

Most students in these categories switch between doing related activities (modeling tools, course MPB: view assignments, see materials, assignment completion, MPB materials) with unrelated activities (others, other assignments and social media). **Students with shorter durations switch across many applications, which explains the negative correlation of duration with the MMT.**

MMT and its Effect on Scores

The correlation analysis indicates that MMT and assignment scores are mostly unrelated. Related activities do not predict scores any better than unrelated activities. This is most likely linked to the nature of the assignment. A1 and A2 instructed the students to create a BPMN model from a certain process description, and they were given approximately two days to submit. No correlation is found between the intensity of MMT with the student's assignment scores, showing that students can still manage the assignment while working with something else.

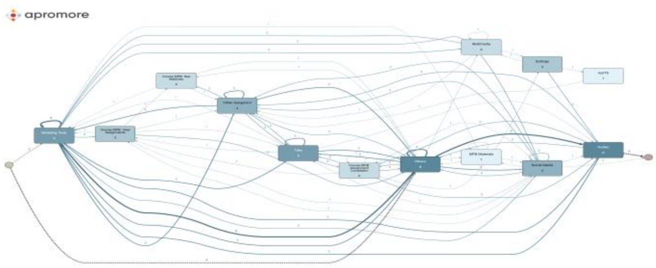

Fig. 3. Students from Batch 1, Class 1 with Less than 2 Hours Duration

The course is designed so students learn and practice during assignments to prepare for the midterm. So, those who are not focused, seen from the high intensity of MMT and prolonged work on the assignments, may find it difficult to do the midterm as it forces them to be more focused within a shorter timeframe. Thus, the lack of statistical significance between students' behavior during assignments, their assignment scores and the midterm results need to be investigated further.

For our analysis, we selected three students (U5 and U24) from Batch 1, Class 2, with relatively low A2 scores compared to their peers. User 24 only started working at midnight before the assignment deadline (**Fig. 4**). For almost six hours, U24 switched between working on the assignment with another assignment and spending more time on unrelated activities than assignment-related activities. The process map of U5 can be found in the supplementary material.

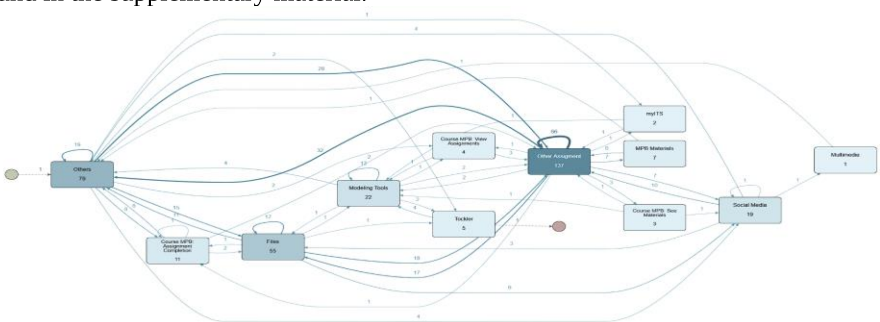

Fig. 4. Process Map of User 24 from Batch 1 while Doing A2 (abstract by case frequency,100 nodes and arcs)

Then, we compared user 11 (U11) from batch 1, which scored very high for A2 and midterm. While doing A2, U11 did more related than unrelated activities, indicating more focus on the assignment (**Fig. 5)**.

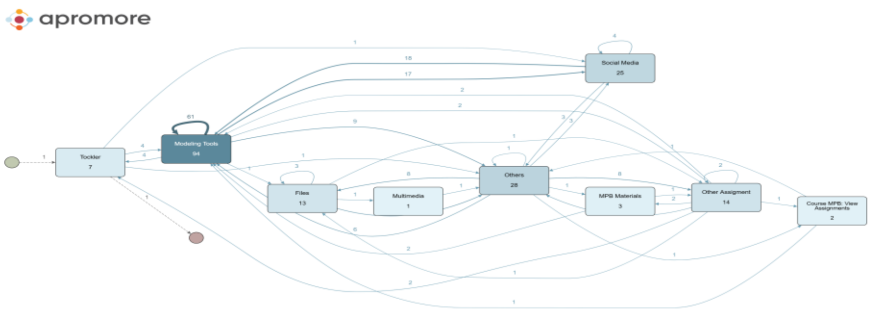

Fig. 5. Process Map of User 11 from Batch 1 while Doing A2 (abstract by case frequency,100 nodes and arcs)

Correlation analysis found the strongest correlation between the midterm's MMT and the number of related activities. Unlike assignments, during midterm, students are only given two hours and cannot open other unrelated applications or windows. **Fig. 6** shows the process map of U9 (Batch 1 Class 2), during midterms that clearly shows the user switch between activities related to the test. **The correlation analysis also found a weak correlation between the duration of the test and the number of unrelated activities**. This is evident in the process map of U30 (Batch 1 Class 2) shown in **Fig. 7**. U30 finished the midterm in 1.88 hours but spent more time with activities unrelated to the midterm. U30 also violated the rules by accessing social media applications and files from the tryout.

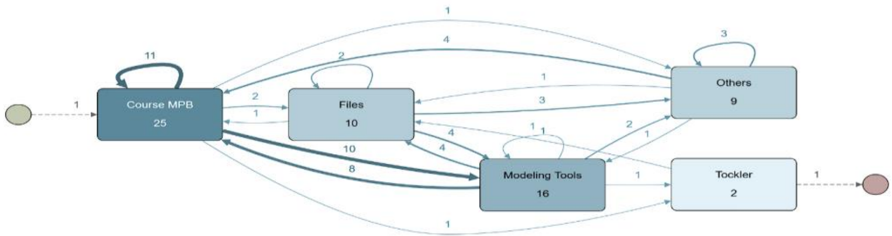

Fig. 6. Process Map of User 9 from Batch 1 During Midterm (abstract by case frequency,100 nodes and arcs)

Submission Time

We also use PM to investigate the students' behavior toward the submission deadline. A strong correlation is found between the time to the deadline of A1 and the hand-in times of A2, indicating that students tend to work similarly for subsequent assignments. We infer from the correlation analysis that students mostly determine their hand-in time based on the deadline, not their schedule. This can be shown in Batch 2, where A1 for class 1 has a different submission time than class 2.

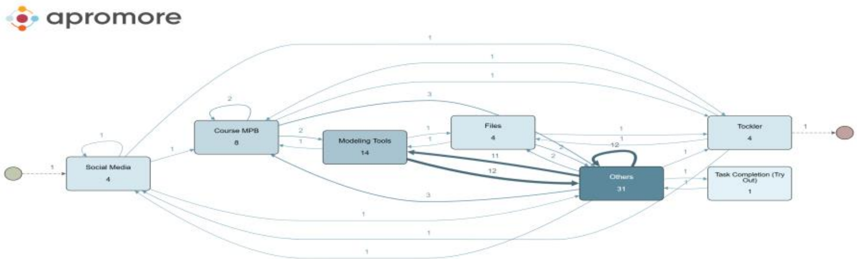

Fig. 7. Process Map of User 30 Batch 1 Class 2 During Midterm (Abstract by case frequency, 100% nodes & arcs)

Fig. 8 shows the timeframe of users from Class 1 in Batch 2 while doing A1. The deadline is 25 March 2024 at 10 PM, and we can see that most students started working on the assignment after midnight before the deadline. **Fig. 9** shows the timeframe of users from Class 2 in Batch 2 while doing A1. For this class, the deadline is 25 March 2024 at 07 AM. Most students started working on the assignment on 24 March at 3 PM until midnight, and some even worked until just two hours before the submission time.

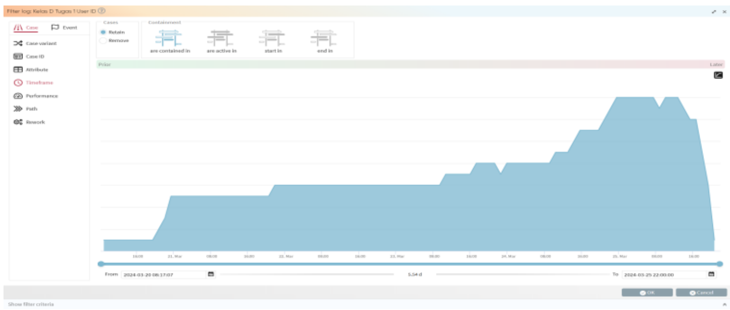

Fig. 8. Timeframe of Users from Class 1, Batch 2 while Doing A1

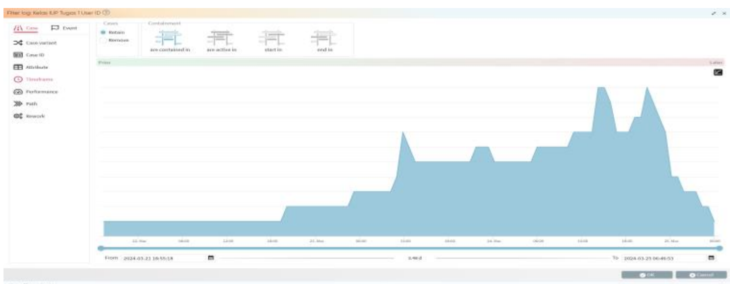

Fig. 9. Timeframe of Users from Class 2, Batch 2 while Doing A1

Finally, we found a moderately strong negative effect between the time to the deadline of A1 and the number of unrelated activities of the midterm. Students who handed in their work closer to the deadline seemed to perform more unrelated activities during the midterm. We chose U2 from Batch 1, Class 1, which handed in A1 at 05.24 AM. The process map of the same student during the midterm is shown in **Fig.10**. U2 conducted many unrelated activities and even violated the rule by accessing social media and files from course material (tryout). This result shows that students who hand in their assignments close to the deadline have lower time management skills and most likely face difficulties during the midterm.

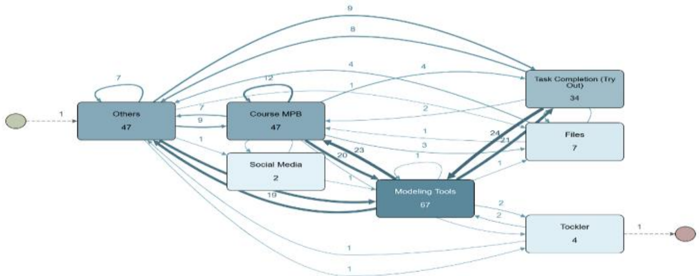

Fig. 10. Process Map of U2, Batch 1, Class 1 during Midterm (abstract by case frequency,100 nodes and arcs)

5 Implications

This paper has two main implications. First, for researchers in PM and EPM, we propose a new way of collecting and processing event logs using Active Window Tracking (AWT). This paper demonstrates that AWT can provide rich insights into students' media multitasking while doing their learning activities. This can be applied to similar digital environments where users perform certain tasks while accessing other media. The recordings can be transformed into event logs and used to investigate the impact of MMT on users' performance.

Second, we provide more practical implications regarding students' MMT behavior and its impact on performance and well-being.

I1. Student MMT and Assignment Performance. Our study found no direct correlation between MMT and students' scores, unlike [14], which reported negative effects. The difference may stem from the assignment type. Process mining shows both high and low scorers multitasked, but low scorers started and finished late, while high scorers focused more on assignment-related tasks.

I2. Student MMT, Last-Minute Habit and Well-being. Correlation analysis indicates a relation of students' hand in time between assignments. PM reveals that students adjust their time of working on the assignment closer to the deadline and often work very late into the night, even until dawn. This highlights their last-minute habit, potentially affecting their well-being, as a study suggests that individuals with an earlier, regular, consistent sleep schedule have better health examination results [10].

Several recommendations can be made to address the issues:
R1. Timed Assignments: To reduce students' last-minute habits, lecturers can set timed assignments, ensuring students have the necessary skills and defining a standard completion time. The submission period should match this standard time.
R2. Adjusted Deadlines: Instead of early morning deadlines, instructors can set submission times during normal working hours to help students manage their time better.
R3. Balanced Workload: Course coordinators and lecturers should consider students' overall workload by coordinating assignment schedules and course deadlines.

Concluding Remark

This paper proposes a new way to collect and process event logs using Active Window Tracking (AWT) to investigate media multitasking (MMT) among students in higher education institutions in Indonesia. Correlation analysis revealed students' consistent behavior across assignments, no direct correlation between MMT and performance, and close to the deadline submission. The PM results revealed that students engaging in MMT frequently switch between assignments, social media, and multimedia. High-scoring students focused more on assignments, while low-scoring students started late, multitasked extensively, and submitted their work close to the deadline. While these results indicate that MMT does not directly affect the performance for the type of assignment, MMT extends work duration. Students tend to work closer to the deadline, so they often work very late into the night, negatively impacting their well-being. Recommendations are provided to mitigate these issues.

Appendix: Supplementary Material

Supplementary material for this article is available online at https://bit.ly/3XUzPPC.

References

1. van der Aalst, W.: Data Science in Action. In: Process Mining. pp. 3–23 Springer Berlin Heidelberg (2016).
2. van der Aalst, W.M.P., Weijters, A.J.M.M.: Process mining: a research agenda. Comput Ind. 53, 3, 231–244 (2004).
3. Alshareef, F. et al.: Educational Data Mining Applications and Techniques. (2020).
4. Bakhshinategh, B. et al.: Educational data mining applications and tasks: A survey of the last 10 years. Educ Inf Technol (Dordr). 23, 1, 537–553 (2018).
5. Beerepoot, I. et al.: A Window of Opportunity: Active Window Tracking for Mining Work Practices.
6. Bogarín, A. et al.: A survey on educational process mining. Wiley Interdiscip Rev Data Min Knowl Discov. 8, (2018).

7. Bradley, V.M.: Learning Management System (LMS) Use with Online Instruction. International Journal of Technology in Education. 4, 1, 68 (2020).
8. Buser, T., Peter, N.: Multitasking. Exp Econ. 15, 4, 641–655 (2012).
9. Cenka, B.A.N. et al.: Analysing student behaviour in a learning management system using a process mining approach. Knowledge Management and E-Learning. 14, 1, 62–80 (2022).
10. Chaput, J.-P. et al.: Sleep timing, sleep consistency, and health in adults: a systematic review. Applied Physiology, Nutrition, and Metabolism. 45, 10 (Suppl. 2), S232–S247 (2020).
11. Cotton, K. et al.: The Effects of Mind-Wandering, Cognitive Load, and Task Engagement on Working Memory Performance in Remote Online Experiments. Exp Psychol. 70, 5, 271–284 (2023).
12. Farrah Siddiqui, I., Ali Arain, Q.: Analyzing Students' Academic Performance through Educational Data Mining. 3C Tecnología. Glosas de innovación aplicadas a la pyme. Edición Especial. 402–421 (2019).
13. Juhaňák, L. et al.: Using process mining to analyze students' quiz-taking behavior patterns in a learning management system. Comput Human Behav. 92, 496–506 (2019).
14. Lerche, T., Kiel, E.: Predicting student achievement in learning management systems by log data analysis. Comput Human Behav. 89, 367–372 (2018).
15. Martín-Perpiñá, M. de las M. et al.: Media multitasking impact in homework, executive function and academic performance in spanish adolescents. Psicothema. 31, 1, 81–87 (2019).
16. Maslovat, D. et al.: Evidence for a response preparation bottleneck during dual-task performance: Effect of a startling acoustic stimulus on the psychological refractory period. Acta Psychol (Amst). 144, 3, 481–487 (2013).
17. May, K.E., Elder, A.D.: Efficient, helpful, or distracting? A literature review of media multitasking in relation to academic performance, (2018).
18. Miller, L.A., Schmidt, J.R.: The Effects of Online Assignments and Weekly Deadlines on Student Outcomes in a Macroeconomics Course. American Economist. 66, 1, 46–60 (2021).
19. Ray, S., Saeed, M.: Applications of educational data mining and learning analytics tools in handling big data in higher education. In: Applications of Big Data Analytics: Trends, Issues, and Challenges. pp. 135–160 Springer International Publishing (2018).
20. Van Der Schuur, W.A. et al.: The consequences of media multitasking for youth: A review, (2015).
21. Uncapher, M.R. et al.: Media multitasking and memory: Differences in working memory and long-term memory. Psychon Bull Rev. 23, 2, 483–490 (2016).
22. Wickens, C.: Attention: Theory, Principles, Models and Applications. Int J Hum Comput Interact. 37, 5, 403–417 (2021).
23. Yadav, C.S., Monga, S.: The Effect of Time Management on Subjective Well-Being Among University Students. International Journal of Management (IJM). 11, 12, 1642–1651 (2020).