Advanced Activity Analysis for Tesys e-Learning platform

Gabriela-Cristiana Costache[1], Paul-Ştefan Popescu[1], Marian Cristian Mihăescu[1], Costel Ionaşcu[2]
[1]Department of Computers and Information Technology,
University of Craiova, Romania
cristianagabriela99@gmail.com, stefan.popescu@edu.ucv.ro, cristian.mihaescu@edu.ucv.ro
[2]Faculty of Economics and Business Administration,
University of Craiova, Romania
icostelm@yahoo.com

ABSTRACT
This paper presents the development and integration of an advanced activity analysis module into the Tesys e-learning platform. Motivated by accreditation requirements and a need for deeper insights into user behavior, the project aimed to provide comprehensive statistics and visualizations of students and faculty engagement. The module was implemented with minimal disruption to existing infrastructure, offering intuitive dashboards, role-based filtering, and data export capabilities. Using anonymized real-world data, the system was evaluated through visual analytics and statistical modeling. Results show significant variance in user engagement and a positive correlation between platform activity and assignment submission. The tool enables evidence-based decision-making for administrators and enhances the overall pedagogical value of the platform.

Author Keywords
Activity Analysys; Human Computer Interaction; E-learning platform; Tesys

ACM Classification Keywords
H.5.m. Information interfaces and presentation (e.g., USI): Miscellaneous.

General Terms
Human Factors; Design; Measurement.

DOI: 10.37789/icusi.2025.18

INTRODUCTION

The goal of this project is to deliver advanced activity analysis capabilities for the Tesys e-learning platform, which supports distance education at the University of Craiova. Currently used by the Faculty of Letters, Tesys enables both students and faculty to access educational resources, complete assessments, and communicate online. However, a key challenge was the lack of accessible data required for meaningful analysis.

To address this, a new administration module was developed to collect and present the necessary information, enabling a clearer understanding of platform usage and user engagement.

The new module aims to provide relevant statistics on user activity (students and professors), including login frequency, participation levels, interaction dynamics, and assessment performance. It will offer administrators a clearer understanding of how the platform is used, thus supporting data-driven decisions for its future development. Another key goal is to ensure seamless integration of the module into the existing system architecture without affecting the core functionalities.

The motivation for the project comes from specific requirements set by ARACIS (the Romanian Agency for Quality Assurance in Higher Education), which mandates tools for monitoring and validating activity on e-learning platforms as part of program accreditation.

As Popescu et. al state, “the interface of Tesys was designed more than 10 years ago and it aimed to fulfill specific needs from that period of time” [1]. Since then, its use has expanded, but it lacks a component to provide an overview of platform activity, now essential in a data-driven digital education environment.

The administrative module seeks to fill this gap by offering accessible, intuitive visualizations of key usage metrics, enabling informed decisions and targeted interventions. Additionally, working with an older, inherited codebase poses technical challenges but also offers valuable learning opportunities and the chance to apply modern solutions to ensure both functionality and compatibility.

RELATED WORK

Understanding user behavior in e-learning platforms has become critical for improving educational outcomes, tailoring content delivery, and ensuring compliance with institutional quality standards. As learning environments shift increasingly online, data-driven insights into student and teacher activity are being used to inform pedagogy, system design, and administrative decisions.

A large body of research has focused on educational data mining (EDM) and learning analytics (LA) to extract actionable information from user interaction data. One of the earliest surveys of EDM techniques applied to virtual learning environments identified clustering, classification, and association rule mining as key tools in understanding learner behavior [2]. Later syntheses noted the increasing use of visual analytics and dashboards to support both learners and instructors in real-time decision-making [3].

Many contemporary learning management systems (LMSs), such as Moodle, Canvas, and Blackboard, incorporate standard analytics features, usually focused on login frequency, assignment submissions, and grades. However, these tools often lack granular tracking, custom visualizations, and institution-specific compliance capabilities [4]. For instance, while dashboards may offer course-level insights, they rarely include role-based filtering or allow bulk data export for external accreditation audits.

Research has also addressed early warning systems and student engagement detection. It has been shown that log data from an LMS can be used to predict student performance, emphasizing the importance of timely and interpretable analytics [5]. Similarly, the development of learning dashboards has been reviewed, with challenges highlighted in balancing usability with analytical depth, particularly when integrating with legacy systems [6].

The topic of educational data mining and student analysis remains a current one, as evidenced by recent research, as stated in [7], which identifies the main perspectives and trends in educational data mining (EDM) within the e-learning environment from a managerial perspective. However, the topic of e-Learning platforms depends on their context and has been debated in many works, such as [8], which analyzes e-Learning platforms based on localization and a sample of 621 students. Still on the localization approach is paper [9] that compares the e-Learning platforms used in the USA with the ones used in Africa. Another analysis refers to the domain of the study, which in the case of [10] is mechanical engineering; in this case, the authors of [10] developed a custom e-learning platform to meet the requirements.

Integrating more technologies to better understand students in e-Learning platforms became a popular task a long time ago, and it is still relevant, as recent studies show, like [11], which reviews research regarding explainable AI approaches in primary education in the context of teaching and learning. Another study that tackles the same problem is [12], which presents a model that achieves 95% precision and generates explanations in less than 500 milliseconds, outperforming traditional methods such as clustering and achieving precision comparable to transformers with significantly lower inference times.

The characteristics of the analyzed studies that explored learning analytics and educational data mining (EDM) from different perspectives are presented below. Their contributions, their limitations, and the contributions of the present paper are contrasted. Romero and Ventura [2] investigated EDM techniques in virtual learning environments (VLE), using clustering and classification methods, but their work was limited by the lack of real-time analytics; our contribution extends this by providing real-time dashboards with role-based access. Papamitsiou and Economides [3] conducted a review of learning analytics literature, identifying as a key limitation the lack of legacy system integration; in contrast, we address this challenge through seamless integration with platforms over ten years old. Macfadyen and Dawson [5] focused on early warning systems through predictive modeling, but their research was restricted to a single-institution context; our work enhances this by offering multi-role analytics along with export capabilities. Finally, Verbert et al. [6] analyzed learning dashboard design in a design review, where limited validation metrics represented a key limitation; we improve on this by implementing comprehensive statistical validation.

Although the existing literature offers extensive coverage of educational data mining techniques and dashboard design principles, it still lacks practical implementations that seamlessly integrate with legacy e-learning systems, provide comprehensive statistical validation with effect sizes and confidence intervals, address specific institutional accreditation requirements and deliver detailed insights into the relationship between platform engagement and academic performance.

This paper addresses these gaps by presenting a complete implementation methodology for legacy system integration, applying rigorous statistical analysis with train/test validation and effect size reporting, explicitly mapping results to ARACIS accreditation requirements, and offering a nuanced analysis that differentiates between engagement probability and completion rates.

PROPOSED APPROACH

To facilitate activity analysis, a new module was added to the platform, providing administrators with a clear and comprehensive overview of user activity. The main objective of this module is to provide the platform administrator with an overview of how students and teachers interact with the platform, based on real data collected and processed in a structured manner.

To achieve this goal, several steps were necessary, including analyzing the existing architecture, identifying relevant data sources, integrating the new functionalities into the current infrastructure, and validating the data displayed in the administrator's interface.

Existing architecture

The architecture of the platform follows a structured model, where each layer is responsible for a specific stage of the application flow. On the frontend, the user interface is built using HTML, CSS, Bootstrap, and JavaScript, integrated into WebMacro (.wm) templates. These templates define page structure and are dynamically populated with data from the backend. User actions (such as selecting a statistic type or submitting a form) trigger server requests.

On the backend, requests are handled by main methods in action classes (each user role has its own action class), following the application’s architecture. These methods interpret the received parameters, validate them, and delegate the logic to an intermediate layer, typically a

manager or service. This layer contains the application logic and interacts directly with the database, executing SQL queries to collect relevant data (e.g., student activity, test results, platform interactions). The retrieved data is then sent back to the WebMacro template to be clearly and interactively displayed to the user.

New module feature integration

The new module includes five main functionalities, each corresponding to a dedicated frontend page and integrated into the navigation bar for easy access.

Students Analysis

Allows viewing and filtering of all registered students by academic and personal criteria (e.g., program, year, name, city). Detailed stats are available per student, number of exams taken, homework completed, messages exchanged, and time spent on the platform. Data can be exported as a .csv file.

This page features a filtering form that allows users to select students based on various general and personal criteria, such as study program, year, name, hometown, workplace, and current status (active/inactive). The form uses Bootstrap components with dropdowns and text inputs for a modern layout.

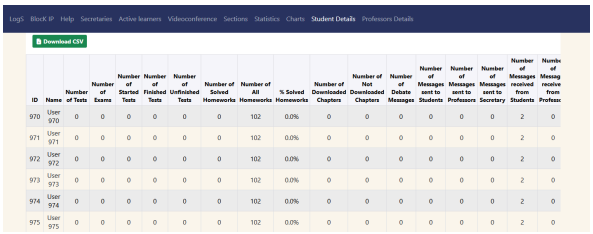

Figure 1. Students details results

The filtered results are displayed in a table, with each row representing a student and showing their calculated statistics. A *Download CSV* button allows exporting the data for further analysis or visualization.

Professors Analysis

Similar to the student section but focused on professors. Filtering includes general academic info and the subject taught. Stats include uploaded chapters, added questions, published homework, messages sent/received, and time on the platform. A .csv export option is also provided.

This page includes a filtering form designed to select teachers based on specific criteria. The interface consists of four main dropdown fields: study program, year, semester, and course. Users can quickly choose predefined values from these menus. The dropdown options are populated by the backend when the page loads. For the courses dropdown, all courses are sent from the backend, and dynamic filtering based on program, year, and semester is handled on the frontend using JavaScript.

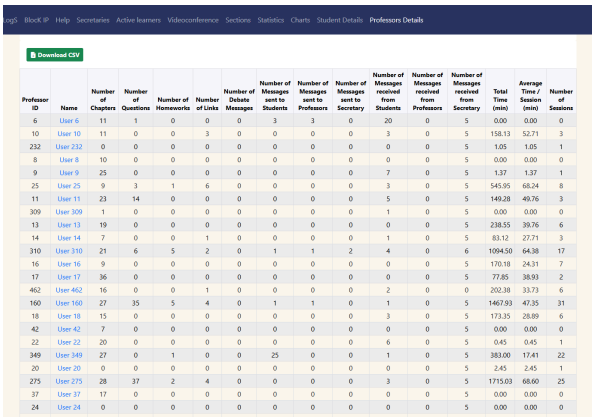

Figure 2. Professors details results

Filtered results are shown in a table, with each row representing a teacher and displaying their calculated statistics. A *Download CSV* button is available for saving the data for later analysis or chart generation.

View Sections

It is used to display all academic programs in a tree structure. Courses are organized by study year and semester, and each course shows related information such as available chapters, assigned homework, and useful links posted by professors.

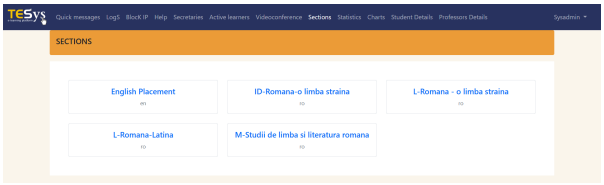

Figure 3. Sections overview

The first version of the section display page includes a representative title followed by a list of Bootstrap cards, each corresponding to a study program. Each card shows the program name and language of instruction. The cards also function as hyperlinks to the next level of view, where courses related to the selected program are displayed.

These courses are organized in an accordion structure, dynamically grouped by year and semester, with sorting handled on the frontend using JavaScript.

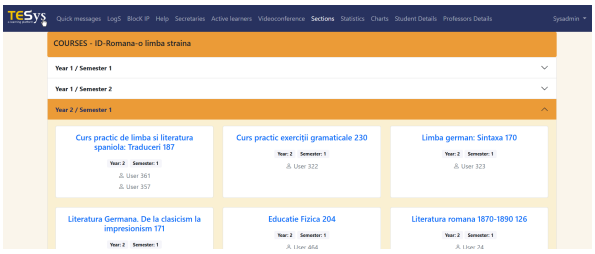

Figure 4. Courses overview

Each course is shown as an informative card that includes the course name, the year and semester it is taught, and the names of the responsible instructors. These cards also act as hyperlinks to the final level of view, where the full list of uploaded materials for each course is displayed. The content is arranged in three separate columns: chapters, homeworks, and useful links, each shown as a list.

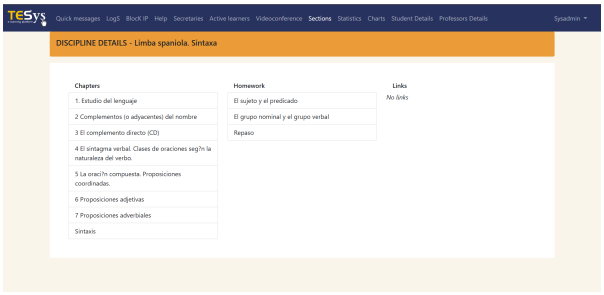

Figure 5. Course details

Charts

Offers dynamic visualizations of platform activity. Users can select a time period to view daily data such as time spent on the platform by students or teachers, downloaded chapters, or homework added by instructors.

This page includes a form with multiple fields that allow users to set a time range using two datepickers: one for the start date and another for the end date. Additionally, the form contains a dropdown to select the type of data to be displayed in the chart (e.g., time spent on the platform, chapters downloaded) and another dropdown to choose the chart style. The visualizations are generated using the Google Charts library.

The charts will be presented in detail in the next section.

General statistics

Provides an overview of platform activity using aggregated statistics, such as total number of students, active students, total chapters, and averages of chapters, homework, and added links per course.

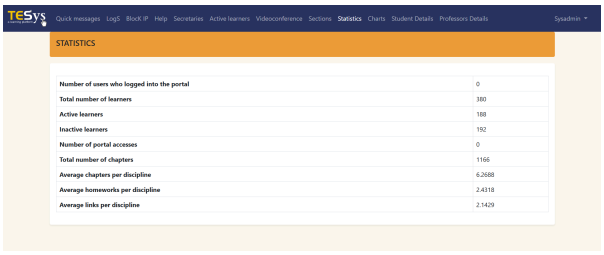

Figure 6. General statistics

EXPERIMENTAL RESULTS

The database used for analyzing the results following the implementation of the new module is a real version of the database used by students and professors of the Faculty of Letters in Craiova. It contains records for activities of 380 students. To comply with legal and ethical requirements regarding the protection of personal data, this database underwent an anonymization procedure, which removed or replaced identifiable information. An initial interpretation of student activity on the platform is provided through the charts available in the *Charts* section. This section features six types of charts that can be seen on Figure 7 a-f).

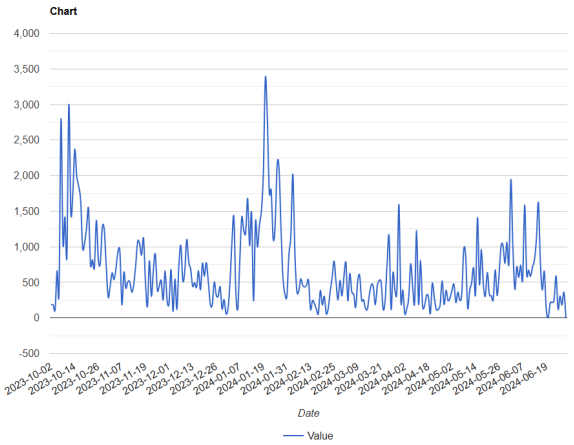

Figure 7. a) Total time spent on the platform by students (in minutes)

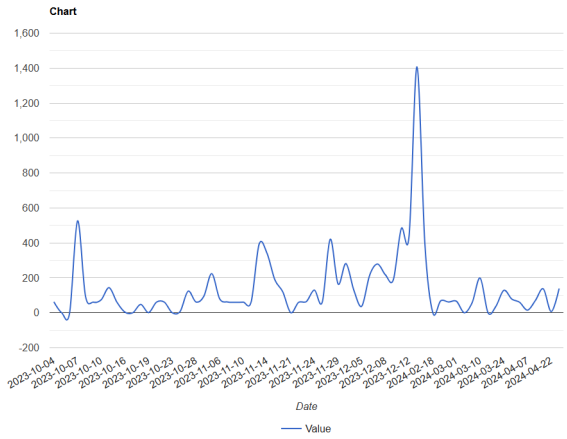

Figure 7. b) Total time spent on the platform by professors (in minutes)

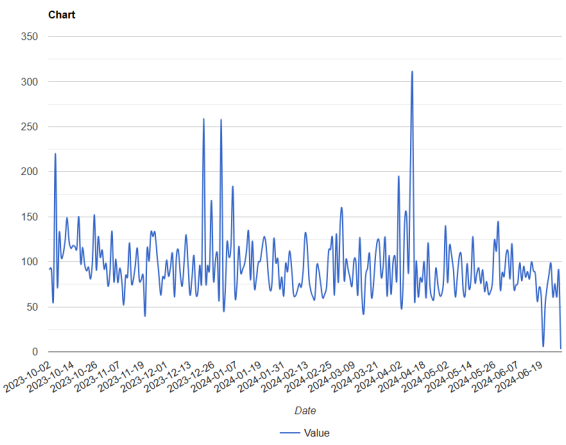

Figure 7. c) Average time spent on the platform by students (in minutes)

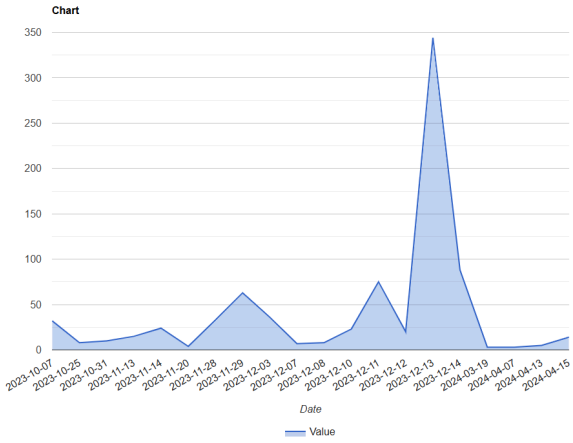

Figure 7. d) Total number of chapters downloaded by students

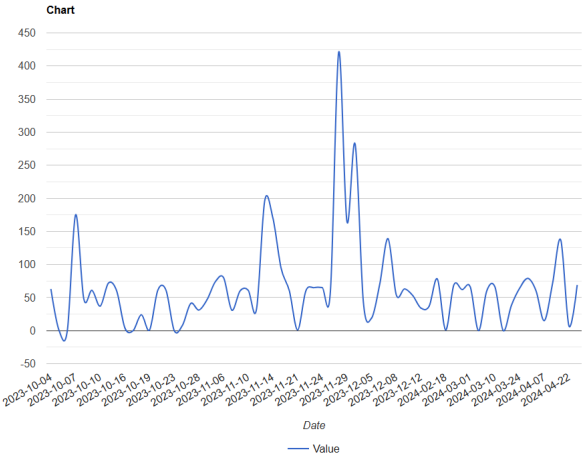

Figure 7. e) Average time spent on the platform by professors (in minutes)

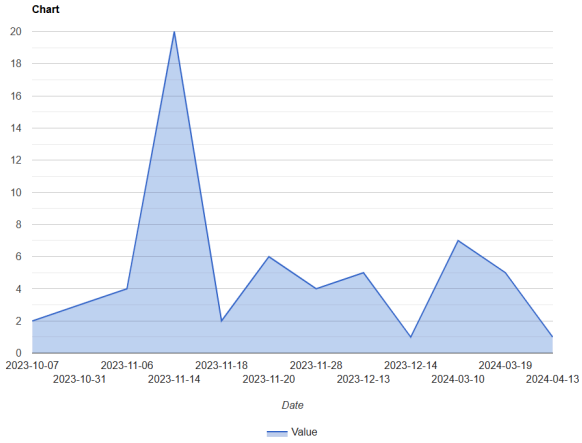

Figure 7. f) Total number of homeworks uploaded by professors

In addition to the data displayed directly on the platform through the newly implemented pages, particularly the section dedicated to charts, further analyses were conducted based on data exported from the system. These were made possible by the table export functionality, which allows users to download information in *.csv* format.

The following section presents an analysis of student-related statistics, based on a dataset containing the following columns:

- ID
- Name
- Number of Tests
- Number of Exams
- Number of Started Tests
- Number of Finished Tests
- Number of Unfinished Tests
- Number of Solved Homeworks
- Number of All Homeworks
- % Solved Homeworks
- Number of Downloaded Chapters
- Number of Not Downloaded Chapters
- Number of Debate Messages
- Number of Messages sent to Students
- Number of Messages sent to Professors
- Number of Messages sent to Secretary
- Number of Messages received from Students
- Number of Messages received from Professors
- Number of Messages received from Secretary
- Total Time (min)
- Average Time / Session (min)
- Number of Sessions

Total Time represents the cumulative duration (in minutes) that each student was connected to the platform during the analyzed period. Average Time/Session is calculated as the ratio between the Total Time and the number of sessions recorded for that user. A session is defined as a continuous period of activity. A session is defined as beginning with login or first interaction and ending at logout or after 30 minutes of inactivity, preventing passive tabs or forgotten logins from inflating time.

We focus on two measures: (1) probability of submitting ≥1 homework (basic engagement) and (2) percentage of homework completion (comprehensive participation). Missing activity metrics were replaced with zeros, assuming absence of logs reflects no platform use during the period.

Some columns in the table contain a very large number of zero values, indicating that many students either did not engage in that type of activity or it was not recorded during the analyzed period. In such cases, interpreting the statistics becomes less relevant or conclusive, as the data distribution is unbalanced, and averages or comparisons may not accurately reflect overall behavior. Variables where more than 80% of observations contained zero values were excluded from correlation and regression analyses, as they provided insufficient variance for meaningful statistical inference. This threshold was established to balance data completeness with analytical validity.

Based on the columns that contain statistically relevant and meaningful data, the following charts were generated.

Distribution of Total Time Spent on the Platform

Using Python's Matplotlib library, a histogram was generated to illustrate the distribution of the total time students spent on the platform, measured in minutes (Figure 8).

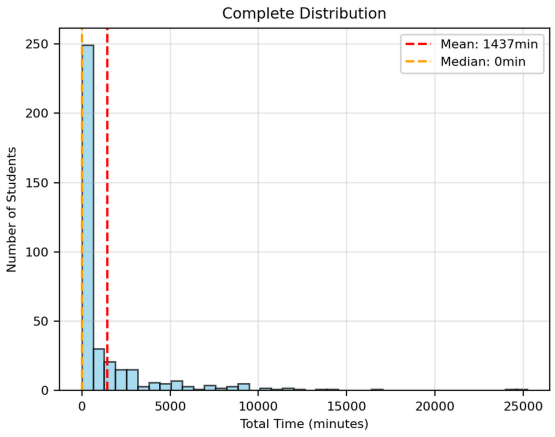

Figure 8. Total time histogram - complete distribution

The chart shows that the overwhelming majority of students logged a low total activity time. There is a strong concentration around lower values, with over 250 students recording minimal connection time.

The distribution of total time is highly right-skewed, with a few students spending thousands of minutes on the platform. The median of 0 minutes, far below the mean of 1,437.37 minutes, together with a coefficient of variation of 216.9%, indicates extreme heterogeneity, where a small group shows intensive use while most display minimal or no activity.

Lorenz Curve for Total Time

The chart in Figure 9 presents the Lorenz Curve for the total time spent on the platform by students and is used to illustrate inequality in activity distribution. The horizontal axis shows the cumulative percentage of students, while the vertical axis represents the cumulative percentage of total time spent online. The dashed orange line represents perfect equality, where all students would have contributed equally to the total time.

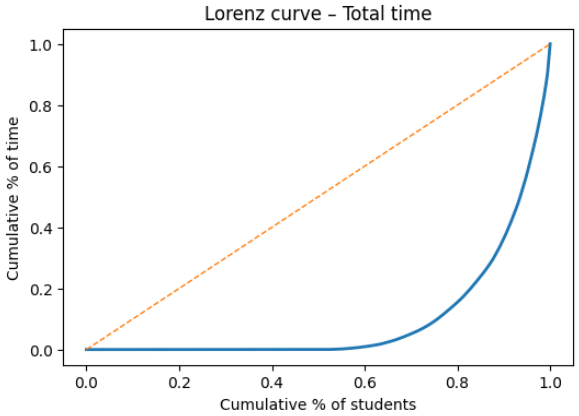

Figure 9. Lorenz curve for total time

The blue curve lies well below the equality line, revealing strong inequality: most students contribute little to total platform time, while the top 10–20% generate nearly all activity. The high Gini coefficient (0.805) confirms this extreme imbalance, underscoring the need for targeted strategies to foster engagement among the majority. This analysis reinforces the histogram's conclusion: platform usage is concentrated among a limited number of users, reflecting an overall imbalance in engagement.

Scatter Plot: Total Time vs. % of Completed Homeworks

The scatter plot in Figure 10 illustrates the relationship between the total time spent on the platform (in minutes) and the percentage of homeworks completed by students. It can be observed that most values for homework completion rates are very low (under 5%), regardless of the total time spent on the platform.

This suggests that spending more time on the platform does not necessarily correlate with a higher rate of assignment

completion. It may indicate passive usage of the platform or difficulties in understanding the tasks.

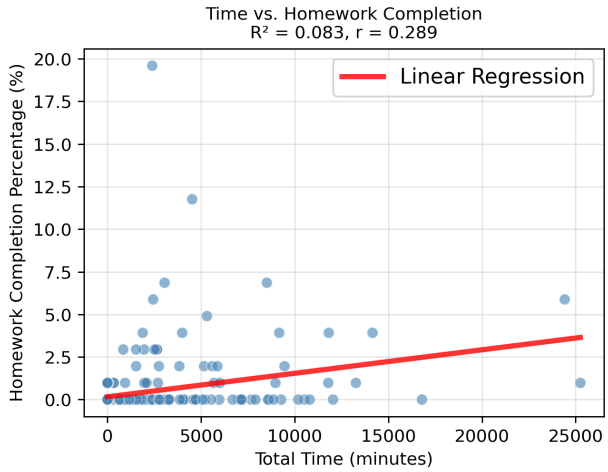

Figure 10. Total time vs. % of completed homeworks

Boxplot of Average Session Time

The chart shown in Figure 11 is a boxplot that illustrates the distribution of the average time students spend in a single session on the platform, measured in minutes. It shows that most sessions are relatively short, with the average close to zero, suggesting that platform interactions are generally brief.

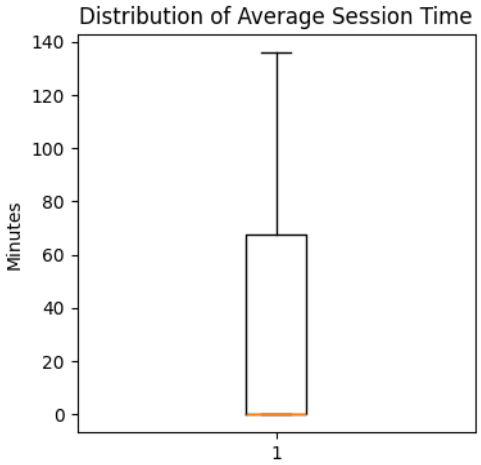

Figure 11. Average session time boxplot

The box represents the interquartile range (between the first and third quartiles), while the whiskers extend to statistically accepted maximum values. A notable outlier (over 130 minutes) indicates the presence of very long sessions, though these are rare. The large gap between the median and the maximum value points to a skewed distribution, with many short sessions and a few significantly longer ones.

Probability of Submitting at Least One Homework

We analyzed the link between total time and homework submission using logistic regression, with data split into training (70%, n = 266) and test (30%, n = 114) sets, stratified to preserve the baseline submission rate (0.132). Results show a significant effect of platform time (β = 0.000220, SE = 0.000052, $p < 0.001$), yielding an odds ratio of 1.000220 (95% CI: [1.000119, 1.000321]). Although the model explains only about 10% of the variance (Pseudo R^2 = 0.0998), this is notable given the complexity of engagement behaviors.

Figure 12. Probability of submitting at least one homework

The chart in Figure 12 shows the relationship between the total time spent on the platform (in minutes) and the probability that a user has submitted at least one homework. The X-axis represents total time, while the Y-axis indicates the probability of having submitted at least one assignment.

Chapters Analysis

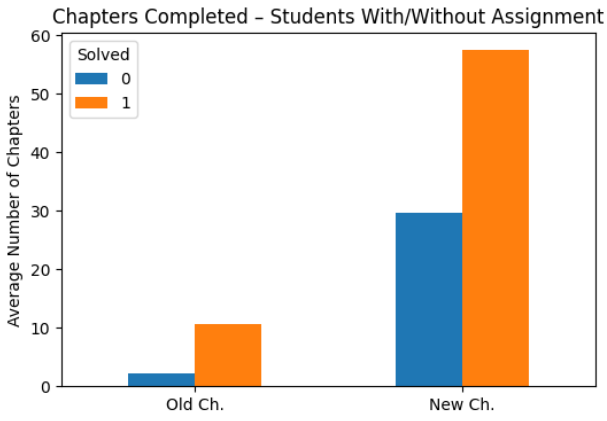

Figure 13. Chapters analysis

The chart in Figure 13 displays the average number of chapters accessed by students who submitted at least one

homework (marked as 1) versus those who submitted none (marked as 0), broken down into two categories: Old Chapters (Old Ch.) and New Chapters (New Ch.). The vertical axis represents the average number of chapters accessed, while the horizontal axis distinguishes the two content types. The *Solved* legend indicates whether a student submitted an assignment or not.

Blue bars represent students who did not submit any homeworks, while orange bars correspond to those who submitted at least one. The data shows that students who submitted homeworks accessed significantly more chapters on average than those who did not. The difference is especially pronounced for new chapters, where the average nearly doubles.

CONCLUSION

The development and integration of the activity analysis module mark a significant step forward in transforming Tesys into a modern, data-informed e-learning platform. By addressing the long-standing lack of visibility into platform usage, the module enables administrators, faculty to better understand how the system is being used and where engagement may be lacking.

The analysis of anonymized, real-use data from the Faculty of Letters revealed several important patterns. First, student activity is highly imbalanced, with the majority spending very little time on the platform while a small minority accounts for most of the engagement. This inequality, confirmed by both histogram and Lorenz curve analysis, suggests that while the platform is functional, its usage is far from uniform, which may reflect differences in motivation, access, course requirements, or digital literacy.

The new module also proved effective from a technical standpoint. Its seamless integration into the legacy system architecture, the use of interactive visualizations, and support for role-based filtering and CSV exports make it both user-friendly and powerful. It aligns with ARACIS requirements for accreditation by providing audit-ready analytics and activity tracking, thereby supporting institutional accountability.

Importantly, the system also sets the stage for future enhancements. With the current framework in place, the platform can evolve to include features such as predictive analytics, early warning systems for at-risk students, and personalized feedback mechanisms. These capabilities are essential in the growing field of learning analytics and critical to improving educational outcomes in digital environments.

Overall, the results highlight that engagement on Tesys is highly uneven, with a small group of students accounting for most of the activity. Since time spent online does not always translate into assignment completion, complementary indicators are needed to better assess learning outcomes. These insights can help administrators support low-engagement students, assist faculty in monitoring participation, and strengthen compliance with ARACIS requirements.

REFERENCES

1. Popescu, P. Ș., Mihăescu, M. C., Mocanu M., & Ionașcu, C. (2016). Evaluation of Tesys e-Learning Platform's Interface. *RoCHI*, 86-90.
2. Romero, C., & Ventura, S. (2007). Educational data mining: A survey from 1995 to 2005. *Expert Systems with Applications*, 33(1), 135–146.
3. Papamitsiou, Z., & Economides, A. A. (2014). Learning analytics and educational data mining in practice: A systematic literature review of empirical evidence. *Educational Technology & Society*, 17(4), 49–64.
4. Ifenthaler, D., Yau, J.YK. (2020). Utilising learning analytics to support study success in higher education: a systematic review. *Education Tech Research Dev 68*, 1961–1990.
5. Macfadyen, L. P., & Dawson, S. (2010). Mining LMS data to develop an 'early warning system' for educators: A proof of concept. *Computers & Education*, 54(2), 588–599.
6. Verbert, K., Govaerts, S., Duval, E., Santos, J. L., Van Assche, F., Parra, G., & Klerkx, J. (2014). Learning dashboards: An overview and future research opportunities. *Personal and Ubiquitous Computing*, 18(6), 1499–1514.
7. Rabelo, A., Rodrigues, M. W., Nobre, C., Isotani, S., & Zárate, L. (2024). Educational data mining and learning analytics: A review of educational management in e-learning. Information Discovery and Delivery, 52(2), 149-163.
8. Wolniak, R., & Stecuła, K. (2024). Evaluation of quality of innovative e-learning in higher education: An Insight from Poland. Applied System Innovation, 7(6), 109.
9. Adeniyi, I. S., Al Hamad, N. M., Adewusi, O. E., Unachukwu, C. C., Osawaru, B., Onyebuchi, C. N., ... & David, I. O. (2024). E-learning platforms in higher education: A comparative review of the USA and Africa. International Journal of Science and Research Archive, 11(1), 1686-1697.
10. Sowndhar, S., Kannan, P., & Sasikumar, P. R. (2024, December). An Experimental Evaluation of Learning Based Platform to Analyze Mechanical Engineering Students Performance. In 2024 International Conference on Innovative Computing, Intelligent Communication and Smart Electrical Systems (ICSES) (pp. 1-4). IEEE.
11. Prentzas, J., & Binopoulou, A. (2025). Explainable Artificial Intelligence Approaches in Primary Education: A Review. *Electronics (2079-9292)*, *14*(11).
12. Villegas-Ch, W., Jaramillo-Alcazar, A., Navarro, A. M., & Mera-Navarrete, A. (2025). Integrating Explainable Artificial Intelligence in Anomaly Detection for Threat Management in E-Commerce Platforms. IEEE Access.