

Data-Driven Analysis of Learning Behaviors Among At-Risk Students Across Disciplines Using Data Mining Techniques

Hon-Sun Chiu, The Hong Kong Polytechnic University, Hong Kong
Adam Wong, The Hong Kong Polytechnic University, Hong Kong
Tung-Lok Wong, The Hong Kong Polytechnic University, Hong Kong

The European Conference on Education 2025
Official Conference Proceedings

Abstract

In recent years, the proliferation of e-learning systems has provided an unprecedented opportunity to collect and analyze vast amounts of data concerning student learning behaviors. This research aims to use data mining techniques to identify common learning patterns of at-risk students across various disciplines. Unlike previous studies that have predominantly focused on individual disciplines or subject areas, this study integrates data from multiple disciplines to offer a comprehensive analysis. We utilized big data sourced from log files of an online e-learning system called Blackboard, supplemented by performance data from individual assessment components. The dataset included all students from 240 subjects across four academic divisions, resulting in a total of 1.7 million rows of records. Our findings show that at-risk students often exhibit low click rates on the e-learning system, delay starting assignments until close to deadlines, and consistently low participation in online activities. Additionally, the study found a high correlation between students' performances in in-class exercises and tests with their overall subject performance. This suggests that regular engagement and participation in formative assessments are strong indicators of academic success. These insights are critical as they will enable educators to develop targeted interventions aimed at improving student performance and retention rates. By understanding the distinctive learning behaviors of at-risk students, educators can provide timely support and resources to help these students succeed. This research not only highlights the importance of using data-driven approaches in education but also underscores the potential of interdisciplinary analysis in enhancing educational outcomes.

Keywords: cross-disciplinary analysis, data mining, at-risk students, e-learning systems, learning analytics

iafor

The International Academic Forum
www.iafor.org

Introduction

The proliferation of digital technologies and the accelerated adoption of e-learning systems in higher education have fundamentally transformed teaching, learning, and the assessment of student performance (Siemens & Long, 2011). In the wake of the COVID-19 pandemic, institutions worldwide rapidly transitioned to online and blended modalities, resulting in a dramatic increase in the generation of rich educational data (Dhawan, 2020). Learning management systems (LMS) now underpin much of the student learning experience, capturing detailed activity logs that chronicle patterns of engagement, resource access, and participation (Romero & Ventura, 2020). These digital footprints present an unprecedented opportunity for educational researchers and practitioners to apply advanced data analytics and artificial intelligence (AI) to support students more effectively, particularly those at risk of academic failure (Baker & Inventado, 2014).

The advent of learning analytics holds significant promise for both educators and learners by enabling the analysis of LMS activity logs in conjunction with assessment records. Such analytical approaches empower educational institutions to transcend generic interventions, facilitating the development of individualized support mechanisms, real-time feedback systems, and targeted mentorship initiatives (Siemens & Long, 2011). Through detailed examination of activity logs within e-learning environments, it becomes possible to identify a range of critical issues, including curriculum planning, monitoring students' learning progress, analyzing learning behavioral patterns, and early identification of at-risk students.

Traditional approaches to identifying at-risk students have relied primarily on lagging indicators, such as grades, attendance records, or subjective educator observations, that often fail to capture the complex, dynamic, and multifaceted nature of student engagement and risk (Arnold & Pistilli, 2012). Such methods may miss subtle early warning signs, resulting in delayed intervention and missed opportunities to prevent academic struggles or attrition (Jayaprakash et al., 2014). Recognizing these limitations, there has been a growing movement toward leveraging educational data mining, machine learning, and predictive analytics to enable early, proactive, and data-driven identification of students who may be at risk across diverse academic contexts (Slater et al., 2017). For example, machine learning can be used to discover patterns from massive student activity data to create a recommendation system for development of generic competencies (So et al., 2023; Wong et al., 2024).

However, the deployment of such systems is not without its challenges. Data quality and integration, ethical and privacy concerns, the imbalance of at-risk versus non-at-risk cases in real-world datasets, and the generalizability of predictive models across subjects and disciplines all pose significant hurdles to the effective and equitable use of AI in education (Baker & Inventado, 2014; Slater et al., 2017). Prior research in the field has often been limited in scope, focusing on single disciplines or individual courses, and utilizing narrowly defined datasets (Jayaprakash et al., 2014). These subject-specific models, while sometimes highly accurate in their own right, frequently lack the capacity to generalize to different contexts where instructional design, assessment structure, and student demographics may vary widely (Romero & Ventura, 2020). As such, there is a pressing need to develop robust, scalable, and interpretable models that can operate effectively across the diversity of modern higher education (Dhawan, 2020).

This research responds to that need by presenting a comprehensive, cross-disciplinary analysis of learning behaviors among at-risk students, utilizing extensive data from the LMS

and assessment records at a higher education institution. By analyzing activity logs and assessment records spanning 240 subjects across four academic divisions, we aim to uncover both universal and context-specific indicators of risk to improve comparability and predictive accuracy across heterogeneous subjects. Critically, the methodology is designed to respect student privacy by excluding sensitive demographic data, focusing solely on behavioral and academic performance metrics available to instructors.

Literature Review

The application of data mining and learning analytics in educational contexts has grown rapidly over the past decade, offering innovative pathways to identify, support, and retain at-risk students (Baker & Inventado, 2014; Romero & Ventura, 2020). Data mining, a subfield of AI, encompasses a broad array of supervised and unsupervised learning algorithms designed to extract meaningful patterns from large, complex datasets (Baker & Inventado, 2014). In educational settings, these methods are increasingly harnessed to analyze students' digital footprints within LMS, digital assessments, and other online platforms (Siemens & Long, 2011). Historically, efforts to identify at-risk students have relied primarily on discipline-specific or course-level models that leverage easily accessible indicators such as prior grades, attendance, assignment submission patterns, and participation in discussion forums (Arnold & Pistilli, 2012; Jayaprakash et al., 2014). These studies often employ classification algorithms to predict student risk based on features tailored to the unique structure of a given course or subject (Arnold & Pistilli, 2012).

Abu-Naser et al. (2015) developed an artificial neural network (ANN) model to predict first-year engineering students' cumulative grade point average (CGPA). Their approach incorporated variables such as academic background, high school characteristics and location, gender, and other demographic data, achieving a prediction accuracy of 84.60%. Similarly, Adekitan and Salau (2019) utilized an ANN model to forecast the academic achievement of engineering students by using the GPA from the first three years to predict the graduation GPA after completing a five-year curriculum, resulting in an accuracy of 89.15%. A recent study demonstrated that utilizing only LMS log data and assessment results can effectively enable the early identification of at-risk students in a computer programming course (Chiu et al., 2023). This approach achieved an accuracy rate of around 85% to 91% in detecting at-risk students in different stages in a semester, as early as the availability of the first assignment's results.

Beyond academic history and prior GPA, additional research has examined the influence of family factors and living environment on student success. For instance, Rivas et al. (2021) considered students' living regions and neighborhood crime rates as predictive variables, while Xu and Chen (2018) included birth dates, number of books borrowed, and internet usage within their models. Hamoud and Humadi (2019) collected data encompassing students' health and sports participation, family and household status, parental circumstances, and social life. Likewise, Rodríguez-Hernández et al. (2021) integrated academic records with socioeconomic status, home characteristics, household information, and employment status to predict student performance. The resulting model accuracies ranging from 78.20% to 87%.

Focusing on learning-specific factors, Saputra et al. (2020) developed a prediction model that achieved a notably high accuracy of 98.20%. Their study aimed to predict student success in an e-learning course, relying primarily on system activity log files for input variables, such as

training sessions, forum participation, chat interactions, discussion contributions, upload assistance, messaging, quiz attempts, and total logins, using gender as the only personal demographic variable.

The literature mentioned above, while effective in producing high accuracy within a constrained context, is inherently limited in generalizability (Jayaprakash et al., 2014). Models trained on data from a single course or discipline may fail to transfer their predictive power to other academic domains with different pedagogical styles, assessment formats, or student populations (Romero & Ventura, 2020). Furthermore, the feature sets selected for these models are often closely tied to the instructional design of specific courses, rendering them less suitable for broader, institution-wide identification of at-risk students (Dhawan, 2020).

There is a pressing need for research that not only integrates diverse sources of behavioral and performance data but also rigorously addresses privacy and ethical considerations, ensuring interpretability and fairness in model deployment (Baker & Inventado, 2014; Dhawan, 2020). This study addresses this gap by employing a comprehensive, cross-disciplinary data mining approach to uncover universal factors among at-risk students. By applying advanced data-driven techniques to a heterogeneous dataset spanning hundreds of subjects and multiple academic divisions, the research advances the field toward scalable, generalizable, and actionable early warning systems in higher education.

Methodology

This research is conducted within the School of Professional Education and Executive Development (SPEED) at The Hong Kong Polytechnic University (PolyU), an institution that provides professional education. PolyU SPEED offers various top-up degree and taught postgraduate programmes across four academic divisions: the Division of Business and Hospitality Management (BHM), the Division of Languages and Communication (LC), the Division of Science, Engineering and Health Studies (SEHS), and the Division of Social Sciences, Humanities and Design (SSHD). There are two main semesters per academic year, each has 13 weeks for teaching and learning activities.

PolyU SPEED uses Blackboard as its LMS for course administration, content distribution, and student engagement. Blackboard allows instructors to develop course sites where students have access to educational resources such as lecture materials and assignments. The system also provides tools for online discussions, assignment submissions, and real-time quizzes to support various teaching practices.

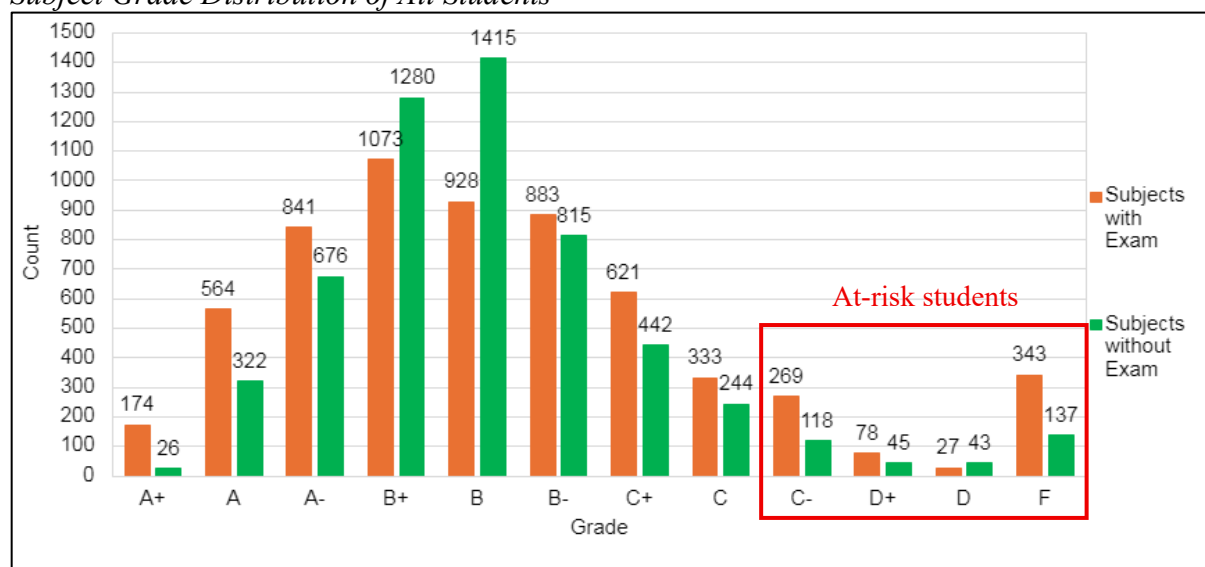
Data Collection

A comprehensive data collection initiative was conducted during the second semester of the 2023–2024 academic year, involving 240 top-up degree subjects across four divisions. This coordinated effort ensured robust representation of the institute's diverse academic programs. Student behavioral data related to learning activities were extracted from the Blackboard, resulting in the compilation of student activity log files, teacher activity log files, and learning component log files. Collectively, these files contained over 1.7 million rows of records documenting various aspects of students' learning behaviors.

Furthermore, comprehensive assessment records for all students were collected and organized by subject within marksheets to support performance analysis. The marksheets contain scores from continuous assessment components and final subject grades for each subject, resulting in over 20,000 assessment records. Figure 1 illustrates the distribution of students across various grades. In this study, at-risk students are defined as those with grades of C- or lower, comprising approximately 5% of the dataset.

Figure 1

Subject Grade Distribution of All Students



The Blackboard log files were systematically summarized, with student activities organized on a weekly basis to accurately reflect learning behaviors over the 13 teaching weeks. Data filtering procedures were applied to ensure reliability. Specifically, subjects with comparable assessment components were selected for detailed analysis. These components included individual assignments, group projects, and tests. Selected subjects encompassed both those with and without final examinations.

This large volume of data not only underscores the scale of the study, but also provides a rich foundation for conducting nuanced analyses that span disciplinary boundaries. The integration of behavioral and assessment datasets supports the development of scalable models for identifying at-risk students, while also enabling the exploration of generalizable factors influencing academic success within a heterogeneous educational environment.

Findings and Discussion

Analysis of the comprehensive dataset revealed a nuanced relationship between student assessment outcomes and their patterns of engagement on the Blackboard e-learning platform. We first analyzed the correlation between assessment components and overall subject grades across the four divisions. Then, the log files extracted from Blackboard are analyzed with the overall subject grades, evaluating the relationship between engagement pattern and academic performance.

Influence of Assessment Components on Academic Performance

Figures 2-7 demonstrate the results of the correlation analysis, where IA and GP stand for individual assignment and group project, respectively. Due to the practical nature of the LC and SSHD, these subjects in Figure 6 and Figure 7 usually do not have examinations.

The correlation results reveal that, as expected, the examination component consistently demonstrates the strongest correlation with overall subject grades among all subjects that incorporate an examination across the four divisions. When the examination component is excluded, which may not be suitable for early detection since it is typically administered at the conclusion of the semester, the test component emerges as the most predictive assessment element, with correlation coefficients frequently surpassing 0.7. Notably, the lowest correlation between the test component and the overall subject grade is observed in BHM subjects that include an examination, as illustrated in Figure 2, where the value registers at 0.59. Nevertheless, this value still represents the highest correlation among all non-examination assessment components under consideration in Figure 2, emphasizing the central role of tests in determining final academic outcomes.

Figure 2
Correlation Analysis of BHM Subjects With Examination

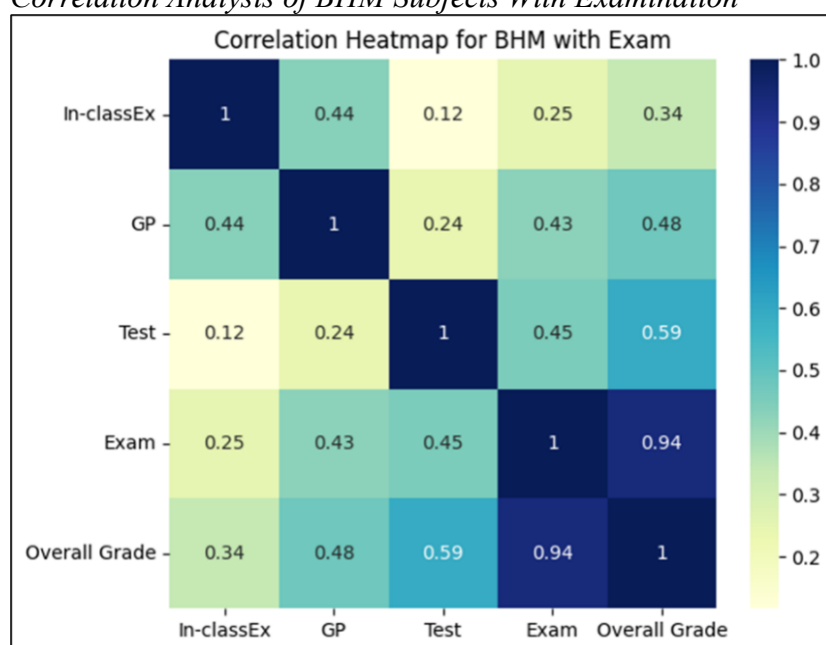


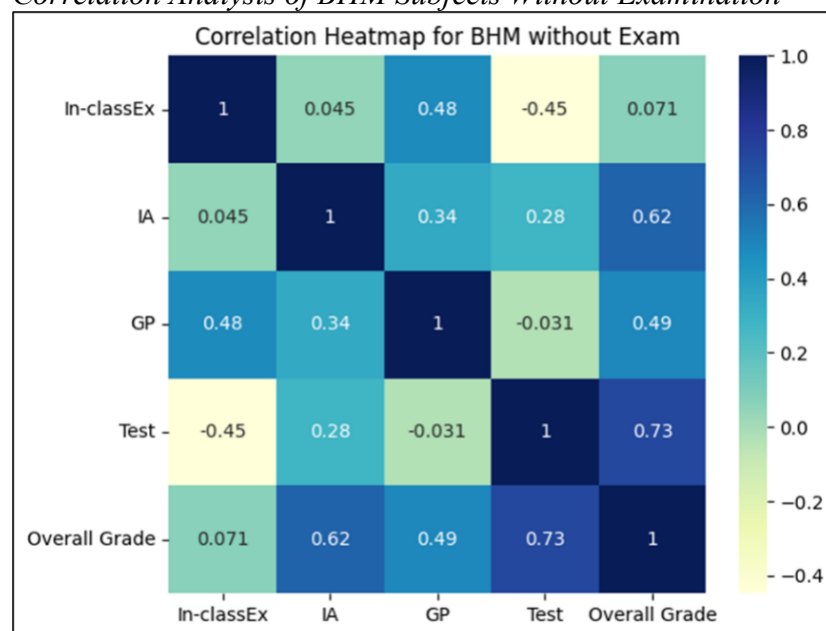
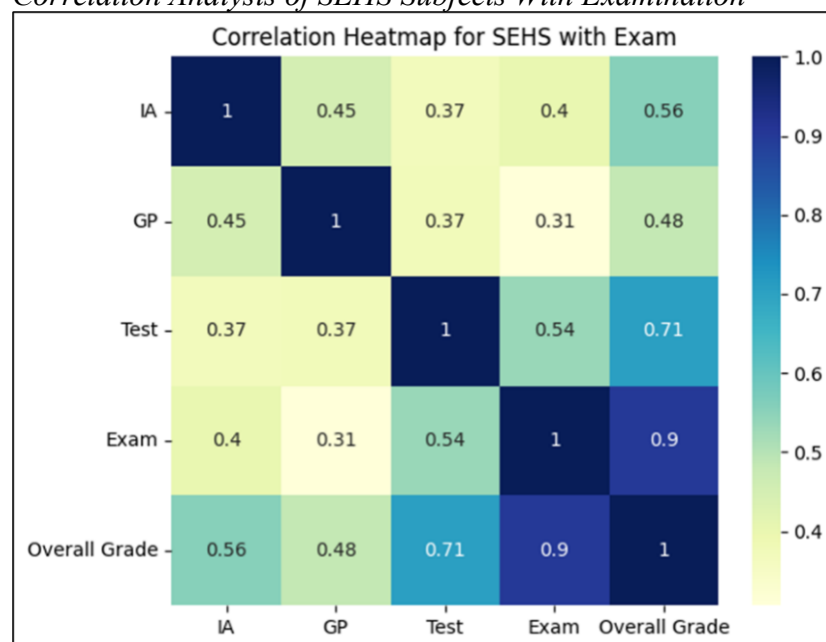
Figure 3*Correlation Analysis of BHM Subjects Without Examination***Figure 4***Correlation Analysis of SEHS Subjects With Examination*

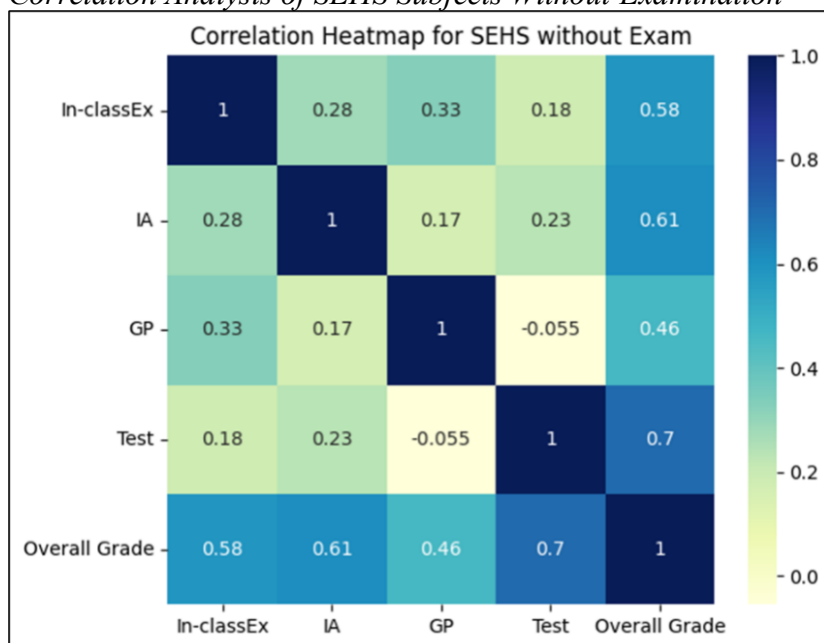
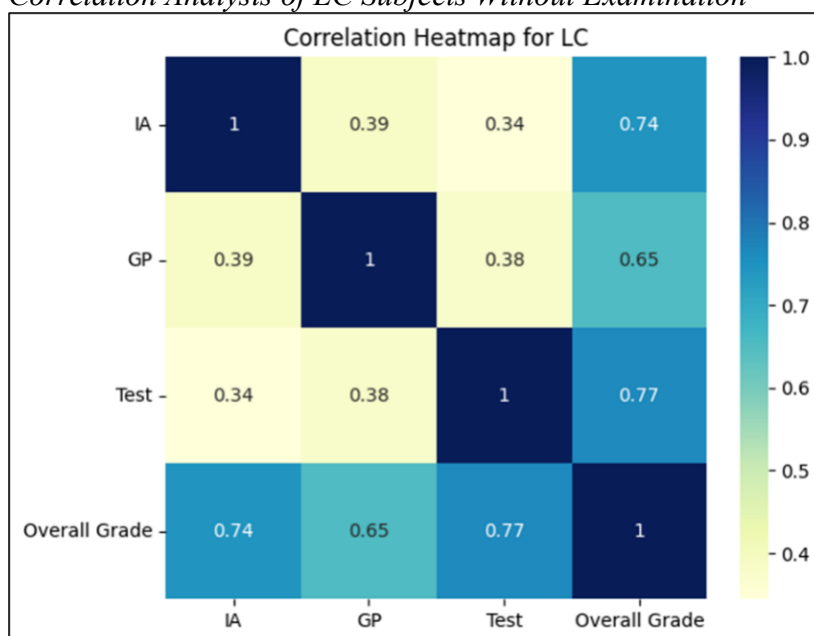
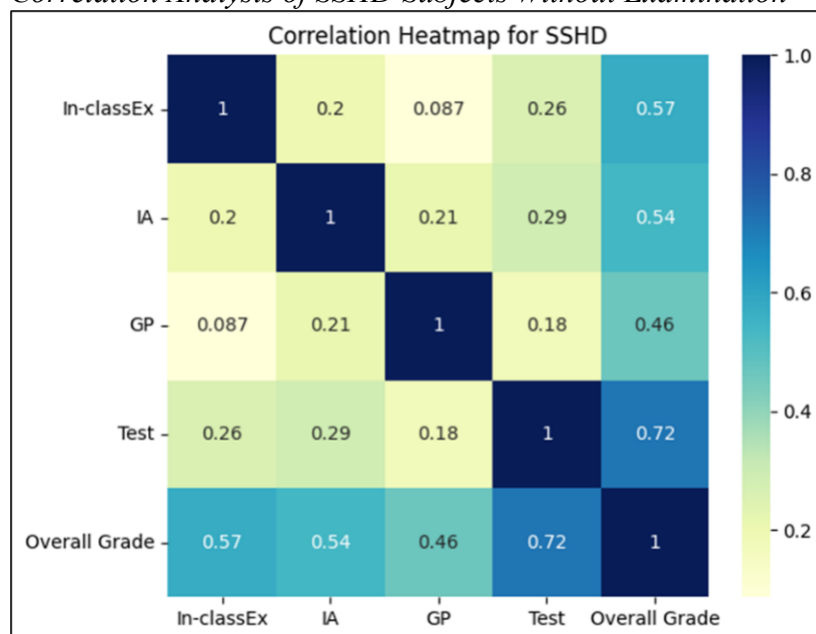
Figure 5*Correlation Analysis of SEHS Subjects Without Examination***Figure 6***Correlation Analysis of LC Subjects Without Examination*

Figure 7*Correlation Analysis of SSHD Subjects Without Examination*

Comparable in nature to tests, typically administered during class time under controlled conditions, in-class exercises similarly require students to complete specific tasks within a predetermined period. To provide a comprehensive analysis, the correlation between in-class exercise performance and overall subject grades was examined wherever such data were available. The findings, however, varied considerably across divisions.

For BHM subjects, as depicted in Figure 2 and Figure 3, the correlation coefficients are notably low, with values of 0.34 and 0.071 for subjects with and without examinations, respectively. These represent the lowest correlations among all assessed components within this division. In contrast, for SEHS subjects, in-class exercises are only present in subjects that do not include an examination. The correlation observed is relatively strong, with a coefficient of 0.58 as shown in Figure 5, a value on par with those for individual assignments and group projects. LC subjects do not incorporate in-class exercises within their assessment frameworks. For SSHD subjects in Figure 7, the correlation between in-class exercise performance and overall grade is measured at 0.57, surpassing the corresponding values for both individual assignments and group projects.

These results carry significant implications for both instructional design and student support mechanisms within multi-disciplinary educational environments. Foremost, the analysis underscores the pivotal role of tests as the strongest predictors of overall academic performance, particularly in subjects where examinations are either central or absent. Consistently high correlation coefficients between test results and final grades highlight that well-constructed tests are not merely evaluative tools but also crucial indicators of student understanding and mastery. The findings also suggest the strategic importance of administering early tests in the semester. Early testing provides instructors with timely and actionable data, enabling them to quickly identify students who may be struggling. This early detection facilitates targeted interventions, tailored support, and adaptive teaching strategies, increasing the likelihood of improved outcomes for at-risk students. By leveraging test results early in the learning cycle, educators can shift from reactive to proactive support, fostering a more equitable and effective academic environment.

In-class exercises, while not universally predictive of final grades, present a more nuanced impact depending on the subject division. For some divisions such as SEHS and SSHD, in-class exercises display moderately strong correlations with academic performance, suggesting these activities can reinforce learning and provide ongoing feedback. However, in other divisions like BHM, the correlation is notably weaker, indicating that the design and integration of in-class tasks may need to be more closely aligned with learning objectives to maximize their impact. The findings also emphasize the importance of context-sensitive use of in-class exercises, which, when effectively incorporated, can supplement formal assessments and enhance the learning process.

Relationship Between LMS Engagement and Academic Performance

While assessment results serve as valuable performance indicators, students' interactions within the LMS can also yield important insights. Through data mining of Blackboard log files, key engagement indicators have been identified, including login frequency and duration, session count, number of clicks, as well as the mean and standard deviation of these metrics. These features demonstrate correlation with academic performance and may facilitate early identification of at-risk students.

Table 1 presents the correlation between key engagement indicators and overall subject grades across the four academic divisions. In contrast to the earlier analysis of assessment components, the examination of Blackboard interaction metrics does not reveal a strong, consistent correlation when all subjects and students are considered collectively. The aggregate relationship between students' assessment outcomes and their digital engagement within the Blackboard platform remains generally weak. This attenuated correlation is largely attributable to the considerable heterogeneity inherent in the pedagogical frameworks and course designs encompassed by the study. Each discipline imposes its own unique expectations regarding online engagement, assignment configuration, and the integration of technology into the learning experience. Such variability substantially diminishes the predictive value of any singular behavioral indicator when generalized across a broad and diverse educational landscape.

Table 1

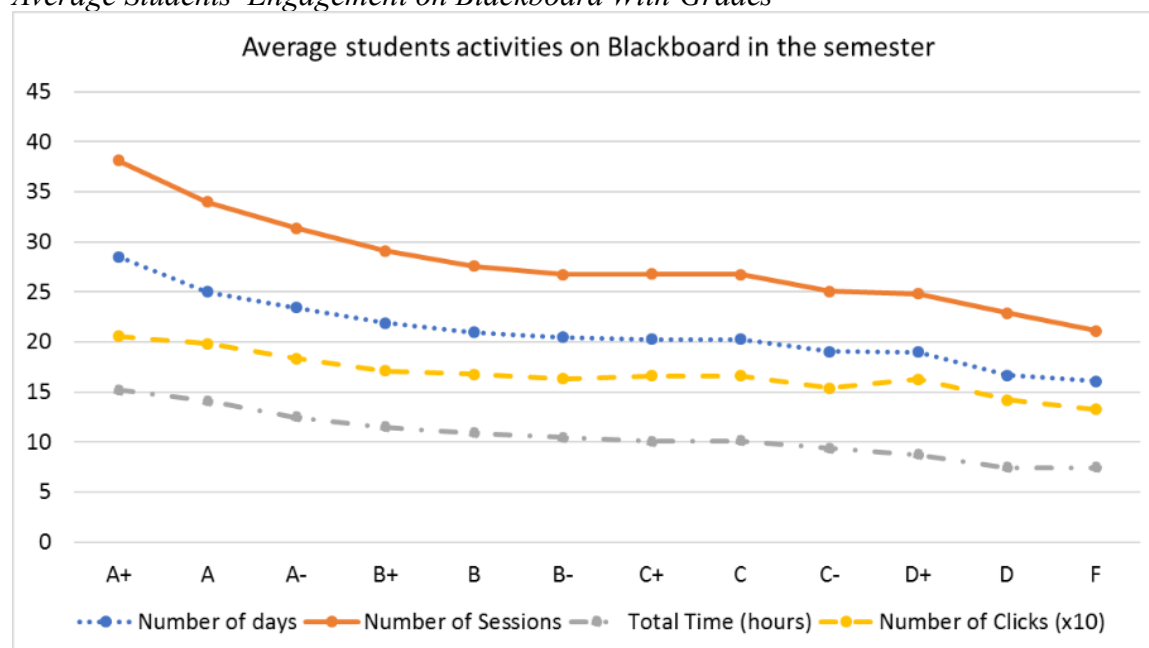
Correlation Between Engagement Indicators and Overall Subject Grade

	BHM	SEHS	LC	SSHD
Days	0.203	0.185	0.142	0.175
Sessions	0.169	0.160	0.116	0.158
Total_Time	0.075	0.123	0.189	0.114
Avg_Total_Time	0.090	0.089	0.154	0.122
Std_Total_Time	0.100	0.118	0.200	0.114
Total_Clicks	-0.064	0.077	0.088	0.030
Avg_Total_Clicks	-0.075	0.036	-0.076	0.077
Std_Total_Clicks	-0.071	0.052	-0.042	0.055
Avg_Time_Per_Day	0.008	0.020	0.108	0.034
Avg_Clicks_Per_Day	-0.213	-0.125	-0.056	-0.192
Avg_Time_Per_Session	-0.008	0.003	0.099	0.018
Avg_Clicks_Per_Session	-0.231	-0.151	-0.081	-0.221

It is particularly noteworthy that the correlation coefficients observed between digital engagement metrics and academic performance in this study are markedly lower than those frequently reported in the existing literature. Previous research, which often focuses on individual courses or homogeneous subject cohorts, has tended to demonstrate moderate to strong associations. The discrepancy observed here is primarily attributable to the unprecedented scale and diversity of the present analysis, which spans 240 distinct subjects, each characterized by individualized instructional approaches, assessment schemes, and activity requirements. The expectations surrounding student engagement with Blackboard vary significantly from one subject to another. For example, dedicating ten hours to activities on the platform may be excessive for a course that emphasizes in-person interaction, yet it might be insufficient for a fully online subject that demands sustained digital participation. Consequently, aggregated measures of engagement lose much of their interpretive power when applied indiscriminately across a highly heterogeneous set of subjects.

Figure 8

Average Students' Engagement on Blackboard With Grades



Despite this overarching limitation, a closer examination of student subgroups reveals a consistent trend. Students with high academic performance, regardless of subject area, tend to exhibit more robust engagement with Blackboard. Figure 8 presents the average students' engagement measures on Blackboard with their overall subject grades. Specifically, good students are characterized by higher click rates on course resources, longer total durations of platform access, and greater login frequency across the teaching weeks under study. These patterns suggest that, while absolute metrics of technology use may not universally predict achievement, proactive and sustained engagement with available digital resources remains a hallmark of successful learners.

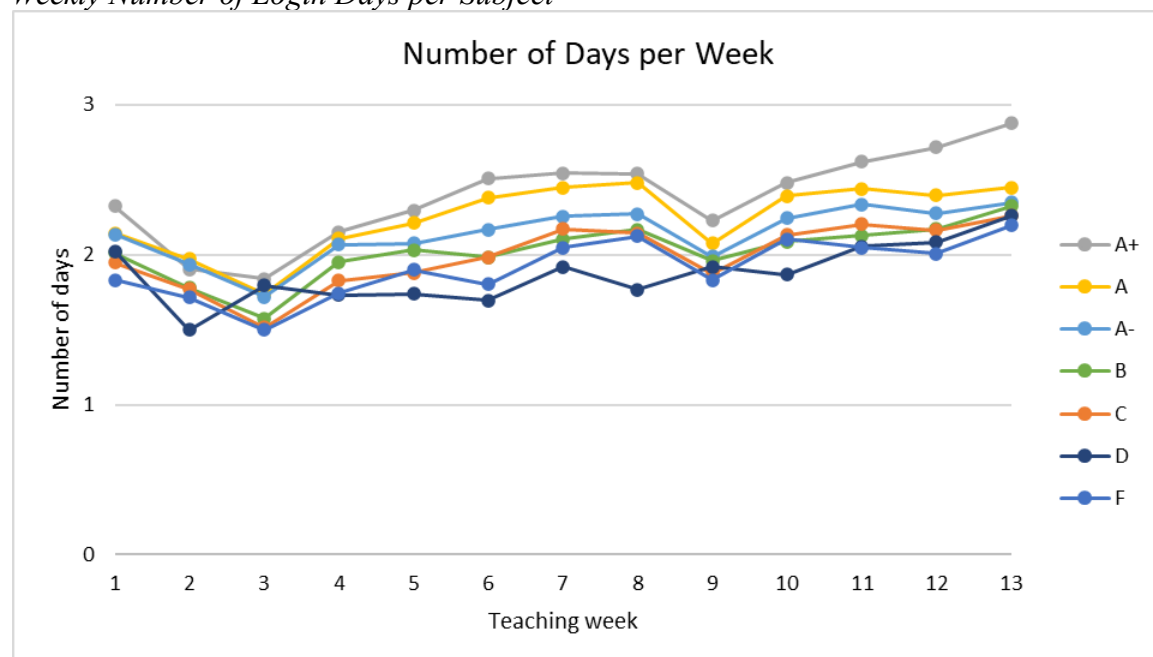
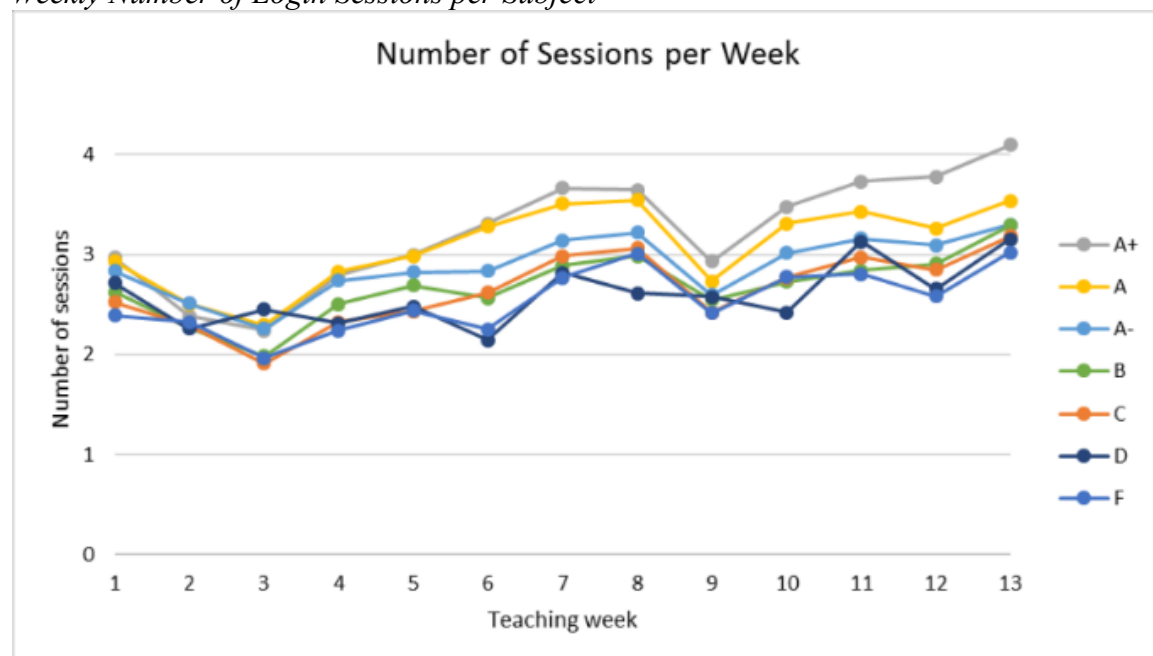
Figure 9*Weekly Number of Login Days per Subject***Figure 10***Weekly Number of Login Sessions per Subject*

Figure 11
Weekly Login Duration (in Hours) per Subject

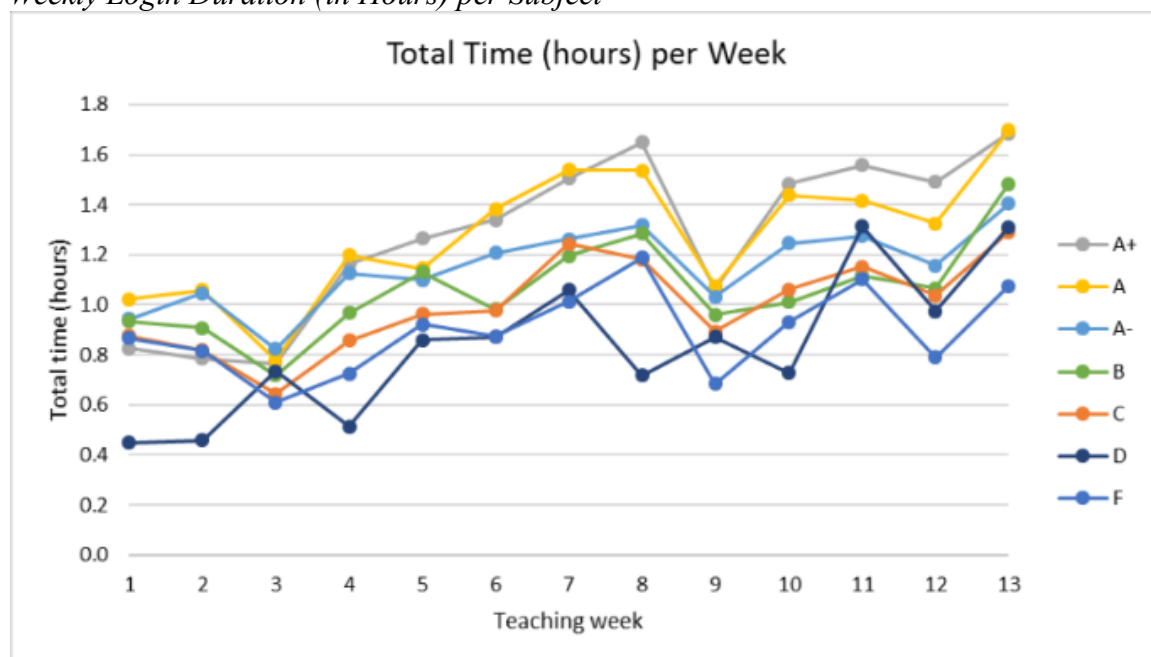
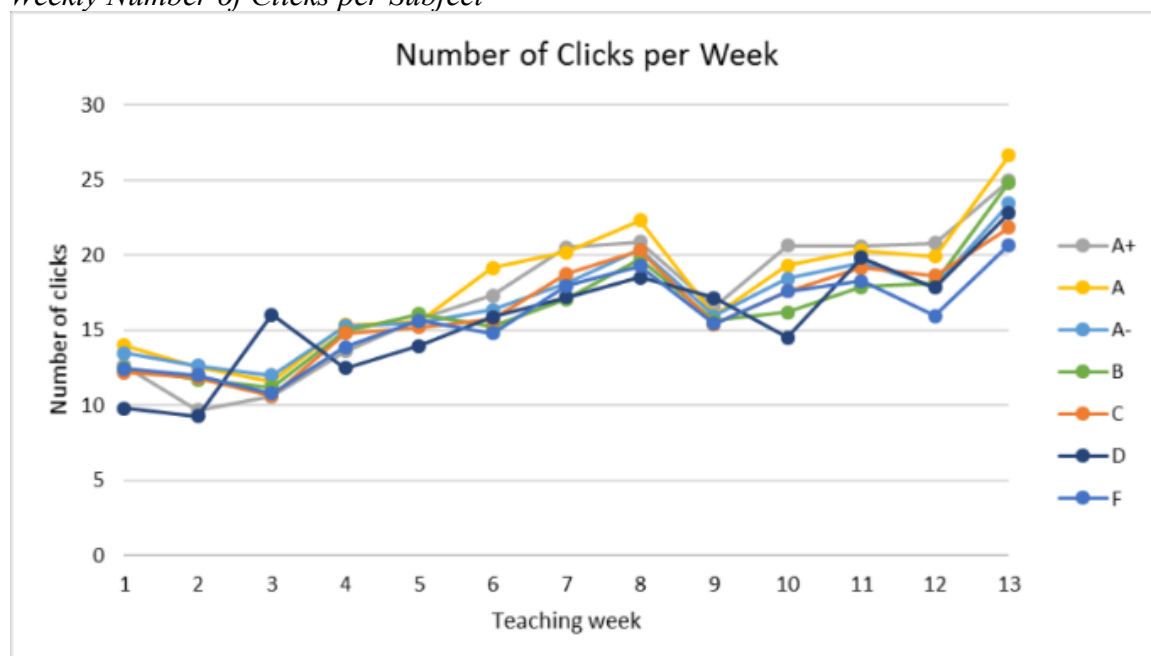


Figure 12
Weekly Number of Clicks per Subject



A more detailed examination of student learning behavior was conducted by analyzing weekly engagement patterns across each subject throughout the 13-week teaching period. As illustrated in Figures 9-12, these patterns were evaluated based on several key metrics: number of days logged in (Figure 9), frequency of login sessions (Figure 10), total login duration per week (Figure 11), and the number of clicks (Figure 12) registered within the subject on a weekly basis.

The findings reveal a general upward trajectory in all engagement metrics among students of varying achievement levels as the academic term progresses. Notably, a consistent and discernible trend emerges: students who attain higher overall grades demonstrate substantially greater engagement with Blackboard. This is evident through their more frequent and sustained interactions, reflected in higher login rates, increased session counts, extended periods of platform access, and a greater number of resources accesses or clicks.

These observations reinforce the notion that, despite the variability in instructional modes and course design, proactive and sustained digital engagement is a characteristic feature of academically successful learners. Such pattern suggests that, while absolute metrics of platform activity may not serve as universal predictors of performance across all contexts, they nonetheless offer valuable insight into the behavioral attributes that distinguish high-achieving students.

Limitations and Future Directions

While the present study draws on a comprehensive and rigorously processed set of digital engagement and assessment data, several limitations must be acknowledged that may shape the interpretation of its findings and inform avenues for future investigation.

First, a fundamental constraint lies in the study's exclusive reliance on data generated within the Blackboard e-learning environment. Although digital behavioral metrics such as login frequency, access duration, and click rates provide valuable windows into patterns of engagement, they inherently omit the substantive domain of students' offline learning behaviors. For instance, after a student opens a file on Blackboard, the dataset cannot discern whether they are actively reading, taking notes, or simply leaving the document open while disengaged. This ambiguity complicates the interpretation of online activity as a direct proxy for cognitive engagement or learning effort. Consequently, some students may appear highly active based on clickstream data, while in reality their true learning may unfold predominantly offline and unobserved by the platform's analytics.

Second, the study does not account for the granularity of assessment design across subjects and instructors. In particular, variations exist in how educators structure and distribute evaluative tasks: some may divide a major assessment into a sequence of smaller quizzes to facilitate continuous monitoring and formative feedback, while others may concentrate evaluation in a single summative exercise. The marksheets collected for this study reflect only the aggregated assessment outcomes and lack visibility into these underlying pedagogical choices and the frequency or sequence of testing events. This omission may obscure nuanced relationships between engagement patterns and achievement, especially where on-going, low-stakes assessments foster more regular student interaction.

These limitations highlight the need for future research to incorporate complementary data sources and methodological innovations. Integrating observational data, self-reported study logs, or in-person classroom analytics could help capture the fuller spectrum of student learning, bridging the gap between online traces and offline effort. Moreover, access to detailed assessment blueprints, including item-level breakdowns and temporal patterns of assignment release, would enable a more precise mapping between instructional design, engagement rhythms, and academic outcomes.

Looking ahead, refining early warning systems and risk identification models will require context-aware analytics that synthesize both digital and offline domains, as well as a deeper appreciation of pedagogical diversity. Expanding partnerships between educators, learning scientists, and data analysts may yield richer, multidimensional datasets, ultimately supporting more equitable and effective interventions for at-risk learners across diverse educational landscapes.

Conclusion

This study underscores the transformative potential of a multidisciplinary, data-driven approach to understanding student learning in contemporary educational environments. By integrating methodologies and insights from educational data mining and learning analytics, the research offers a nuanced examination of how digital engagement and assessment structures collectively shape academic outcomes.

Two principal findings emerge from this analysis. First, the composition and sequencing of assessment components exerts a significant influence on subject performance. Courses that incorporate early-stage tests and well-structured in-class exercises are better positioned to identify students who may be at risk, thereby enabling timely and targeted interventions. Such formative assessment strategies not only provide valuable feedback for learners but also foster more continuous and meaningful engagement with course material. Second, the study reveals a robust and consistent relationship between students' engagement with the Learning Management System (LMS) and their overall academic success. High-achieving students exhibit markedly greater frequency and depth of interaction with the LMS, as reflected in higher login rates, longer access durations, and more resource interactions. This pattern highlights the importance of promoting regular and active use of digital learning platforms as a vehicle for reinforcing academic habits and supporting sustained achievement.

The implications of these findings are multifaceted. Educational practitioners are encouraged to embed early diagnostic assessments and scaffolded in-class activities within their curricula, not only to enhance learning outcomes but also to create mechanisms for early risk detection. Simultaneously, fostering a culture of frequent and purposeful engagement with the LMS can serve as a catalyst for improved student performance. Future initiatives should continue to bridge disciplinary boundaries, leveraging the combined strengths of technology, pedagogy, and analytics to support more personalized, equitable, and effective learning experiences for all students.

Acknowledgement

The work described in this paper was fully supported by a grant from the Research Grants Council of the Hong Kong Special Administrative Region, China (Project No., UGC/FDS24/H06/23).

References

- Abu-Naser, S., Zaqout, I., Abu-Ghosh, M., Atallah, R., & Alajrami, E. (2015). Predicting Student Performance Using Artificial Neural Network: In the Faculty of Engineering and Information Technology. *International Journal of Hybrid Information Technology*, 8(2), 221–228. <https://doi.org/10.14257/ijhit.2015.8.2.20>
- Adekitan, A. I., & Salau, O. (2019). The impact of engineering students' performance in the first three years on their graduation result using educational data mining. *Heliyon*, 5(2), Article e01250. <https://doi.org/10.1016/j.heliyon.2019.e01250>
- Arnold, K. E., & Pistilli, M. D. (2012). Course Signals at Purdue: Using learning analytics to increase student success. *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge* (pp. 267–270).
- Baker, R. S. J. d., & Inventado, P. S. (2014). Educational data mining and learning analytics. In J. A. Larusson & B. White (Eds.), *Learning analytics: From research to practice* (pp. 61–75). Springer.
- Chiu, H. S., Ng, C. T. B., & Cheung, Y. L. (2023). Academic Performance Prediction using Artificial Neural Networks based on Course-related Student Data. *International Journal of Advance Computational Engineering and Networking (IJACEN)*, 11(8).
- Dhawan, S. (2020). Online Learning: A Panacea in the Time of COVID-19 Crisis. *Journal of Educational Technology Systems*, 49(1), 5–22.
- Hamoud, A., & Humadi, A. (2019). Student's Success Prediction Model Based on Artificial Neural Networks (ANN) and A Combination of Feature Selection Methods. *Journal of Southwest Jiaotong University*, 54(3). <https://doi.org/10.35741/issn.0258-2724.54.3.25>
- Jayaprakash, S. M., Moody, E. W., Lauría, E. J. M., Regan, J. R., & Baron, J. D. (2014). Early alert of academically at-risk students: An open-source analytics initiative. *Journal of Learning Analytics*, 1(1), 6–47.
- Rivas, A., González-Briones, A., Hernández, G., Prieto, J., & Chamoso, P. (2021). Artificial neural network analysis of the academic performance of students in virtual learning environments. *Neurocomputing*, 423, 713–720. <https://doi.org/10.1016/j.neucom.2020.02.125>
- Rodríguez-Hernández, C. F., Musso, M., Kyndt, E., & Cascallar, E. (2021). Artificial neural networks in academic performance prediction: Systematic implementation and predictor evaluation. *Computers and Education: Artificial Intelligence*, 2, Article 100018. <https://doi.org/10.1016/j.caeai.2021.100018>
- Romero, C., & Ventura, S. (2020). Educational data mining and learning analytics: An updated survey. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 10(3), e1355. <https://doi.org/10.1002/widm.1355>

- Saputra, E. P., Supriatiningsih, Indriyanti, & Sugiono. (2020). Prediction of Evaluation Result of E-learning Success Based on Student Activity Logs with Selection of Neural Network Attributes Base on PSO. *Journal of Physics: Conference Series*, 1641, 012074. <https://doi.org/10.1088/1742-6596/1641/1/012074>
- Siemens, G., & Long, P. (2011). Penetrating the fog: Analytics in learning and education. *EDUCAUSE Review*, 46(5), 30–40. <https://doi.org/10.17471/2499-4324/195>
- Slater, S., Joksimovic, S., Kovanovic, V., Baker, R. S., & Gasevic, D. (2017). Tools for educational data mining: A review. *Journal of Educational and Behavioral Statistics*, 42(1), 85–106. <https://doi.org/10.3102/1076998616666808>
- So, C., Wong, K. L., Tsang, K. H. Y., Chan, A. P. L., Wong, S. C. W., & Chan, H. C. (2023). Some pattern recognitions for a recommendation framework for higher education students' generic competence development using machine learning. *Journal of Technology and Science Education*, 13(1), 104–115. <https://doi.org/10.3926/jotse.1707>
- Wong, S., Wong, K. L., Lau, Y. Y., Tsang, K., & Chan, A. (2024). Perceived Usefulness of a Machine Learning-Assisted Recommendation System for Generic Competency Development. *Journal of Education and e-Learning Research*, 11(3), 614–621. <https://doi.org/10.20448/jeelr.v11i3.5971>
- Xu, W. W. M., & Chen, X. (2018). A Performance Predictor for Honors Students Based on Elman Neural Network. *ITS 2018 Workshop Proceedings*. p. 55.

Contact email: honsun.chiu@cpce-polyu.edu.hk