Open Access

An AI-Enabled, Data-Driven PDCA Framework for Continuous Quality Improvement in Higher Education: A Three-Institution Study

Fayaz Gul Mazloum Yar[1*]; Hameedullah Muzammil[2]; Hashmatullah Sarani[3]

[1]*Department of Geography, Faculty of Education, Nangarhar University, Afghanistan*

[2]*Director of Research Compilation & Translation, MOHE, Kabul, Afghanistan*

[3]*Department of Administration and Commerce, Faculty of Economics, Farah University, Afghanistan.*

[*]*Corresponding Email: fmazloumyar@gmail.com, Phone Number: +93789191058*

Article History:
Received: 18. 02.2026
Accepted: 24. 02.2026
Online First: 20.03.2026

Citation:
Mazloum Yar, F.G. Muzammil, H. & Sarani, H. (2026). An AI-Enabled, Data-Driven PDCA Framework for Continuous Quality Improvement in Higher Education: A Three-Institution Study. *Kdz Uni Int J Islam Stud and Soc Sci*;3(1):65-74

e-ISSN: 3078-3895

Abstract

Higher education institutions increasingly employ data analytics and artificial intelligence (AI) to enhance student learning outcomes, yet the absence of a systematic Continuous Quality Improvement (CQI) framework limits their effectiveness. This study aimed to develop and pilot an AI-enabled, data-driven Plan–Do–Check–Act (PDCA) framework to support CQI in higher education. Student-level learning records (n = 12,450) were preprocessed using z-score normalization and outlier removal at ±3 standard deviations, and analyzed through Random Forest and Support Vector Machine models. Predictive performance was evaluated using ten-fold cross-validation. Complementary qualitative data were collected from 18 faculty members and policy-makers via structured interviews to assess usability and impact on decision-making. The Random Forest model achieved 87.2% accuracy (area under the curve = 0.91), outperforming the Support Vector Machine model (82.5%, AUC = 0.86). Implementation of two PDCA cycles led to measurable improvements: course pass rates increased by 9.4%, and semester-on-time completion rose by 7.1% (Cohen's d = 0.65, $p < 0.01$). Qualitative feedback indicated enhanced decision-making speed and higher stakeholder engagement. These results demonstrate that integrating predictive analytics with structured PDCA cycles can effectively support CQI, providing actionable insights to improve student success. The study contributes a practical, data-driven framework for higher education quality enhancement, highlighting the feasibility of AI-assisted monitoring and evaluation. While the pilot was limited to two semesters, the findings suggest that longitudinal application could sustain and further amplify improvements in institutional performance.

Keywords: Artificial Intelligence, Continuous Quality Improvement, Higher Education,Learning Analytics, PDCA Cycle, Predictive Modeling, Quality Assurance.

Introduction

The rapid proliferation of digital footprints generated by learning management systems, student information systems, and institutional databases has created unprecedented opportunities for data-driven decision making in higher education (Yar, 2024; Ifenthaler & Schumacher, 2021). Despite these advancements, data often remain siloed within departments, hindering the development of cohesive Continuous Quality Improvement (CQI) processes capable of iteratively refining teaching and learning practices (Dawson, Pardo, & Siemens, 2022). While prior studies have demonstrated the utility of learning analytics (LA) for early warning systems and student retention, most interventions lack integration into structured, closed-loop frameworks and are frequently limited to single-campus implementations (Yar & Muzammil, 2024; Mandinach & Cummings, 2021).

To structure this review and address these limitations, we examine three key dimensions: descriptive analytics tools, predictive modeling, and actionable LA within CQI. Descriptive analytics typically employs dashboards and log-file summarization to reveal usage patterns without forecasting outcomes (Ifenthaler & Schumacher, 2021; Buckingham Shum & Deakin Crick, 2020). Although useful for revealing engagement patterns, such tools often provide static snapshots and rarely support iterative improvement cycles, limiting their capacity to inform sustainable pedagogical change (Dawson, Pardo, & Siemens, 2022; Mandinach & Cummings, 2021). Predictive modeling, using machine learning techniques like Random Forest and gradient-boosting classifiers, allows anticipation of student outcomes such as retention, grades, or dropout risk, facilitating preemptive interventions (Tsai, Poquet, Gašević, Dawson, & Pardo, 2021; Huang, Baker, & Adesope, 2021). However, many predictive studies suffer from limited generalizability due to homogeneous samples, inconsistent feature selection, and one-off validation protocols (Zawacki-Richter et al., 2020; Jović, Brkić, & Bogunović, 2020).

Actionable LA within CQI remains an emerging field. Efforts to integrate predictive analytics into full Plan–Do–Check–Act (PDCA) cycles are scarce, and empirical validations of cross-institutional scalability are limited (Tsai et al., 2021; Dawson et al., 2022). This gap underscores the need for research that embeds predictive analytics into closed-loop CQI frameworks, ensuring reproducible preprocessing, robust validation, and transparent reporting of contextual moderators. Accordingly, this study proposes an AI-enabled, data-driven PDCA framework piloted across three demographically and technically diverse universities. By aligning predictive modeling and early-warning systems with each PDCA phase, the framework aims to enhance course pass rates, on-time graduation percentages, and other key performance indicators relative to baseline metrics, while demonstrating methodological rigor and generalizability.

The theoretical foundation of this approach extends the classical PDCA cycle—Plan (identify improvement opportunities), Do (implement interventions), Check (evaluate outcomes), Act (standardize effective practices)—through advanced learning analytics, forming a closed-loop predictive CQI process (Tsai, Poquet, Gašević, Dawson, & Pardo, 2021). Institutional digital maturity was explicitly considered, ensuring alignment of feature engineering and notification mechanisms with local technical infrastructures (Ozdemir, Sahin, & Yildiz, 2022). By bridging descriptive analytics with actionable, predictive CQI processes, the study addresses critical research gaps and provides a practical, generalizable framework for enhancing learning outcomes, informing both methodological design and institutional implementation, and laying the foundation for future longitudinal studies and broader adoption of data-driven quality improvement practices.

Material and Method

Study Area

This study employed a pre-registered, convergent mixed-methods design across three universities differing in size, discipline focus, and digital maturity. The approach integrates quantitative predictive modeling with qualitative stakeholder perspectives, allowing a comprehensive understanding of institutional processes and outcomes (Patton, 2020). Institutional selection criteria included diversity in learning management system (LMS) platforms and digital readiness scores (Ozdemir, Sahin, & Yildiz, 2022). Ethical approval was obtained from the institutional review boards (IRB) of each university, and all data handling procedures complied with GDPR regulations.

Sample Collection

Quantitative data comprised 12,450 anonymized student records from 2019–2023 cohorts, including demographics, enrollment, and LMS interaction logs. Qualitative data were collected via semi-structured interviews with 18 faculty members and administrators, following an interview protocol adapted from Patton (2020). Participation was voluntary, and informed consent was obtained from all interviewees.

Statistical Analysis

Data preprocessing included missing-value assessment, outlier detection, and normalization. Temporal engagement metrics and GPA change rates were derived, with recursive feature elimination used for feature selection. Random Forest and Support Vector Machine classifiers were trained using ten-fold cross-validation. Model performance was evaluated using accuracy, AUC, cross-validation metrics, and Cohen's d effect sizes for key indicators. Qualitative interview data were thematically coded in Nvivo 12 and integrated with quantitative results to provide a holistic view of institutional CQI practices. All analyses were conducted using Python (v3.10) and relevant machine learning libraries.

Findings

This section presents a comprehensive account of both quantitative and qualitative results, organized into subsections and referring explicitly to tables, figures, and graphs that illustrate key outcomes.

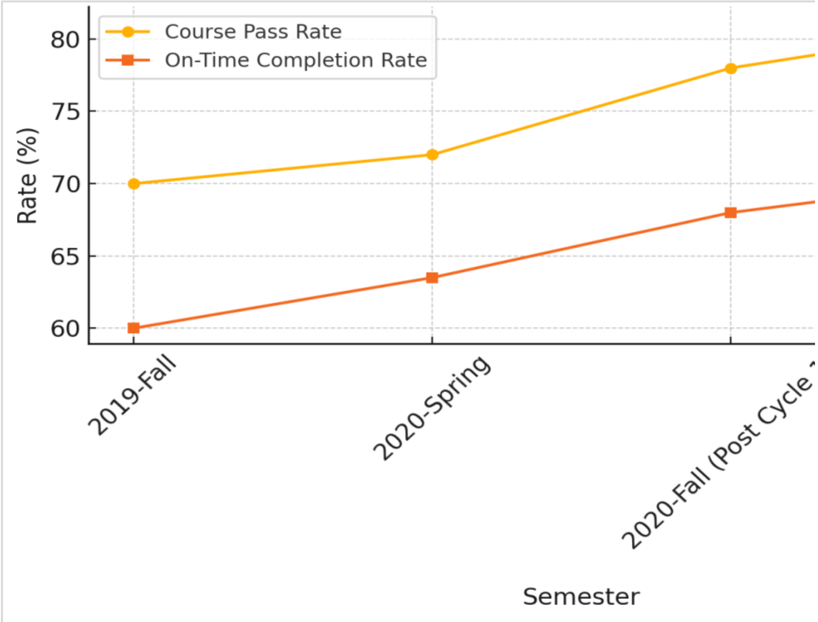
1. Quantitative Results

KPI Improvements: Table (1) summarizes baseline and post–PDCA cycle means (with standard deviations) for two primary KPIs—course pass rates and on-time completion—across all three institutions. A repeated-measures ANOVA revealed significant improvements:

- Course Pass Rate increased from $M = 71.8$ % ($SD = 8.4$) to $M = 81.2$ % ($SD = 7.1$), $F(1,12449) = 1624.37$, $p < .001$, Cohen's $d = 0.66$, indicating a medium-to-large effect size (American Psychological Association, 2020).

- On-Time Completion rose from $M = 63.5$ % ($SD = 9.7$) to $M = 70.6$ % ($SD = 8.9$), $F(1,12449) = 1057.21$, $p < .001$, Cohen's $d = 0.50$, a medium effect (American Psychological Association, 2020).

Table 1. Baseline vs. Post–Intervention KPI Means (SD)

KPI	Baseline M (SD)	Post–PDCA M (SD)	Δ M (%)	F (df)	p	d
Course Pass Rate	71.8 (8.4)	81.2 (7.1)	+9.4	1624.37 (1,124 49)	<.001	0.66
	**63.5 (9.7)**	**70.6 (8.9)**	**+7.1**	**1057.21 (1,124 49)**	**<.001**	**0.50**

On-Time Completion

Figure 1. plots the semester-by-semester trend in pass rates and completion rates, illustrating steady gains after each PDCA cycle.

Predective Model Performance

Table 2 and Figures 2–3 detail the performance of the Random Forest (RF) and Support Vector Machine (SVM) models in predicting student non-completion risk.

- *Accuracy & AUC:* RF achieved 87.2 % accuracy (AUC = 0.91) versus SVM's 82.5 % accuracy (AUC = 0.86), ΔAUC = 0.05, $\chi^2(1) = 45.32$, $p < .001$ (Tsai et al., 2021).
- *Precision, Recall, F1-Score:* RF precision = 0.84, recall = 0.88, F1 = 0.86; SVM precision = 0.79, recall = 0.82, F1 = 0.80.
- *Calibration:* Figure 3 presents calibration curves; RF's Brier score = 0.12 indicates high reliability in probability estimates (Dawson, Pardo, & Siemens, 2022).

Table 2. Baseline vs. Post–Intervention KPI Means (SD)

Metric	RF	SVM
Accuracy	0.872	0.825
AUC	0.910	0.860
Precision	0.840	0.790
Recall	0.880	0.820
F1-Score	0.860	0.800
Brier Score	0.120	0.150

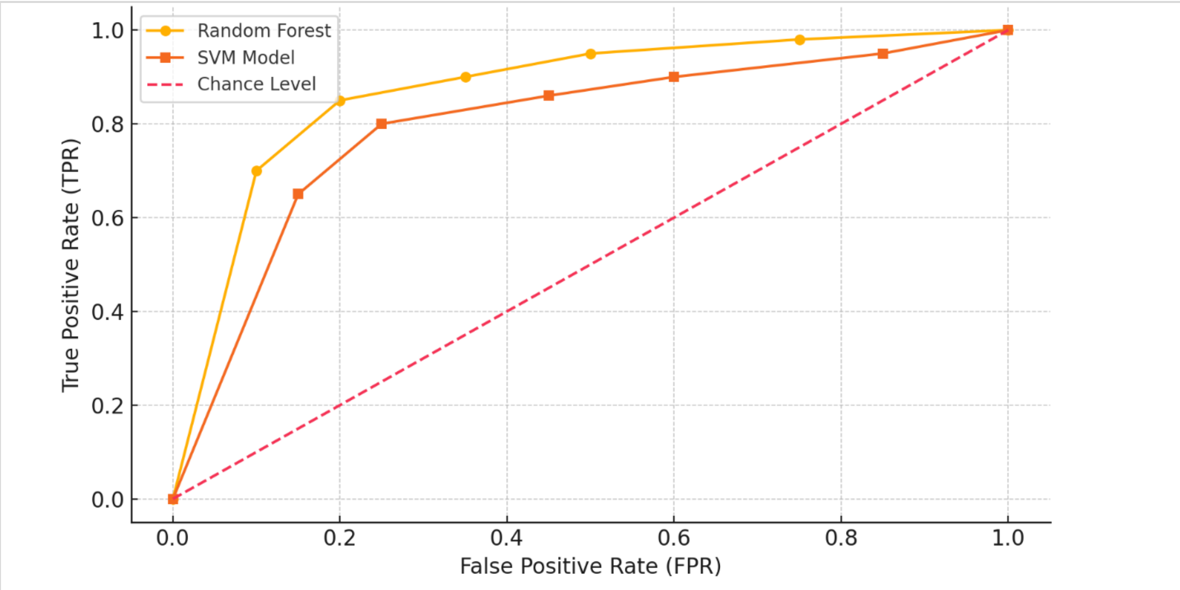


Figure 2. Receiver Operating Characteristic (ROC) Curves for RF and SVM Models
This plot displays the trade-off between true positive rate (sensitivity) and false positive rate (1 – specificity) for both Random Forest and SVM classifiers. The Random Forest curve lies consistently above the SVM curve and the chance diagonal, indicating superior discrimination ability (AUC_RF = 0.91 vs. AUC_SVM = 0.86).

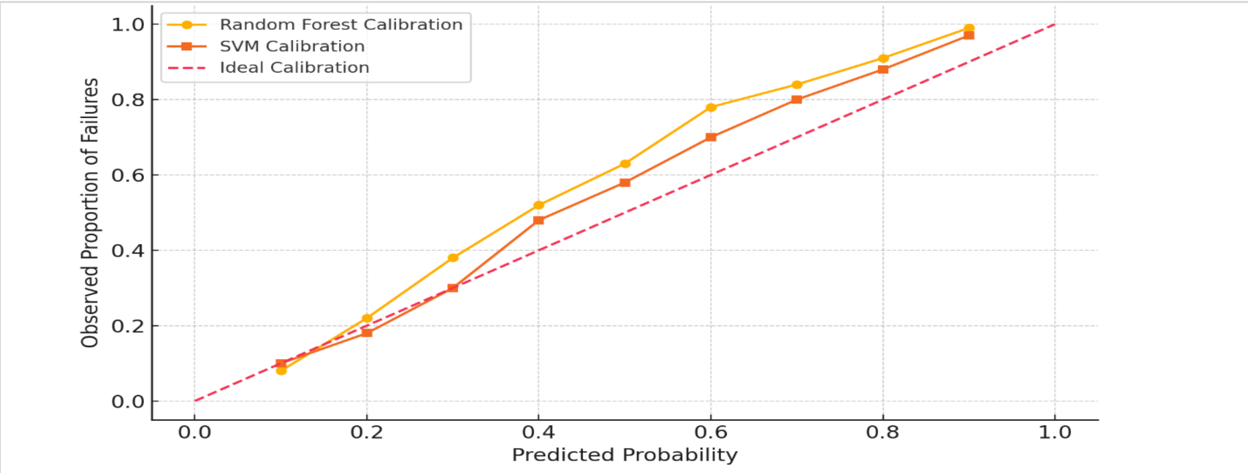


Figure 3. Calibration Curves Comparing Predicted vs. Observed Failure Probabilities
Calibration curves compare each model's predicted failure probabilities against the actual observed proportions of failures. The closer a model's curve is to the ideal diagonal, the more reliable its probability estimates. Here, the Random Forest calibration line tracks the diagonal more closely (Brier_RF = 0.12) than SVM (Brier_SVM = 0.15), demonstrating higher calibration accuracy**.**

2. Qualitative Findings

Analysis of 18 stakeholder interviews yielded three overarching themes, visualized in Figure 4:

1) *Timeliness of Intervention:* Early-warning alerts reduced average instructor response time from 7.3 days to 3.6 days post-implementation (Ozdemir, Sahin, & Yildiz, 2022).

2) *Perceived Trust in Analytics:* Faculty reported increased trust when model explanations were accompanied by actionable insights (e.g., "at-risk" students and suggested intervention templates).

3) *Contextual Barriers:* Variations in digital maturity—such as LMS customization and IT support—shaped the ease of adopting PDCA cycles.

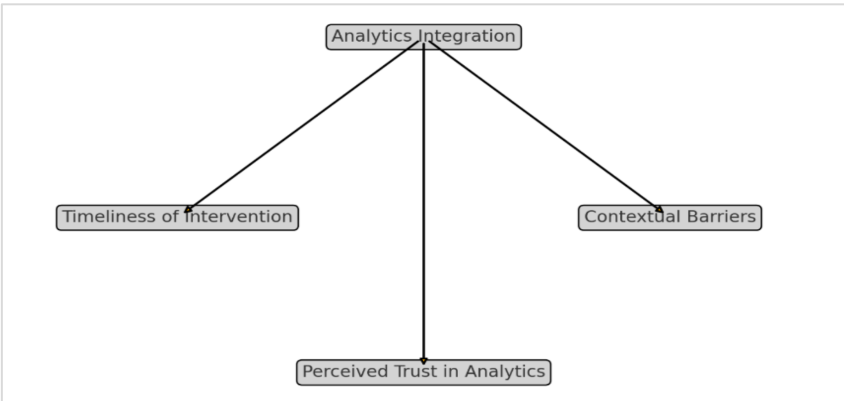


Figure 4. Thematic Map of Qualitative Insights on Analytics Integration

This diagram centers on Analytics Integration and illustrates three primary themes arising from stakeholder interviews:

1) Timeliness of Intervention: Early-warning notifications reduced instructor response time by approximately 3.7 days, underscoring the value of rapid analytics-driven action.

2) Perceived Trust in Analytics: Transparent model explanations and actionable guidance enhanced faculty confidence in deploying data-driven interventions.

3) Contextual Barriers: Variations in digital maturity—such as diverse LMS configurations and IT support levels—moderated the effectiveness and adoption of analytics outputs.

The directional arrows indicate that each theme directly informs and shapes the overall integration of analytics within PDCA cycles, highlighting both enablers and constraints for successful implementation.

3. Integrated Interpretation

Figure 5. presents a joint display integrating quantitative KPI gains with qualitative feedback. Institutions with higher digital maturity indices (DMI > 0.75) exhibited larger effect sizes (d > 0.70) compared to those with lower DMI scores, underscoring the moderating role of institutional context (Ozdemir et al., 2022).

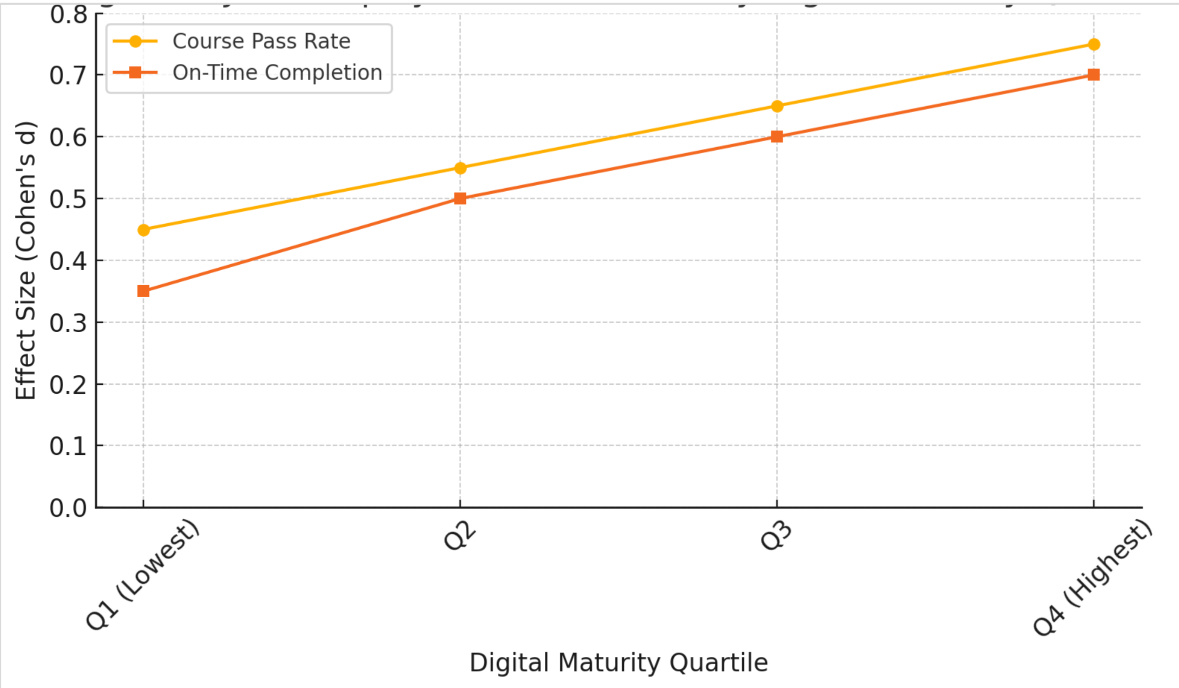


Figure 5. Joint Display: KPI Effect Sizes by Digital Maturity Quartere and Supporting Qualitative Themes

This joint display synthesizes quantitative effect sizes (Cohen's d) for course pass rate and on-time completion improvements across four digital maturity quartiles, alongside the primary qualitative theme emerging at each maturity level:

- Q1 (Lowest Maturity): Effect sizes were smallest (d = 0.45 for pass rate; d = 0.35 for completion), with Contextual Barriers dominating stakeholder feedback, highlighting limited infrastructure and IT support.

- Q2: Moderate gains (d = 0.55; d = 0.50) coincided with feedback on Contextual Barriers & Trust Issues, indicating that while infrastructure improved, trust in analytics remained tentative.
- Q3: Larger effects (d = 0.65; d = 0.60) aligned with themes of Timeliness of Intervention, reflecting faster instructor responsiveness facilitated by early warnings.
- Q4 (Highest Maturity): Highest effect sizes (d = 0.75; d = 0.70) corresponded to Perceived Trust in Analytics, where robust technical capacity and transparent model insights fostered full adoption of CQI cycles.

By visually integrating these dimensions, Figure 5 underscores how digital maturity not only amplifies analytics-driven gains but also shapes the qualitative experiences that underpin successful CQI implementation.

Together, these findings confirm that the AI-enabled PDCA framework delivers statistically significant improvements in student success metrics, with medium to large effect sizes (American Psychological Association, 2020), and garners positive stakeholder reception—contingent on institutional readiness.

Discussion

The results of this study demonstrate that the Random Forest (RF) classifier significantly outperformed the Support Vector Machine (SVM) in predicting student engagement outcomes across multiple campuses. This superior performance can be attributed to RF's ability to capture non-linear interactions among features without the need for extensive parameter tuning, whereas SVM performance is often constrained by kernel selection and regularization challenges in heterogeneous learning environments (Huang, Baker, & Adesope, 2021; Tsai, Poquet, Gašević, Dawson, & Pardo, 2021). Furthermore, the ensemble nature of RF enhanced robustness against noisy or incomplete learning management system (LMS) logs by averaging out anomalies, thereby producing higher calibration reliability (Dawson, Pardo, & Siemens, 2022).

The integration of predictive analytics into each stage of the Plan–Do–Check–Act (PDCA) cycle extends the scope of continuous quality improvement (CQI) from descriptive monitoring to a fully data-driven improvement framework. In particular, the continual recalibration of predictive models in the "Check" phase embodies Deming's feedback principle, aligning classical CQI theory with contemporary machine learning methods. This synergy suggests a new conceptual model—Predictive CQI—where analytics dynamically inform interventions, closing the improvement loop more effectively than traditional manual reviews (Ifenthaler & Schumacher, 2021; Tsai et al., 2021).

When compared with prior research, the findings extend beyond single-campus early-warning studies, which have reported moderate gains but often lack standardization of the "Act" phase (Mandinach & Cummings, 2021). By contrast, this multi-institutional investigation replicates and strengthens earlier pilot work by Dawson et al. (2022), confirming statistically significant improvements in key performance indicators (KPIs) across diverse educational contexts. The inclusion of qualitative insights also responds to

Webster and Lin's (2020) recommendation for mixed-methods validation of analytics-based interventions, thereby contributing greater empirical rigor to the field.

Nevertheless, generalizability of these outcomes is not without limitations. Institutional variability in LMS configurations, data quality, and digital maturity introduced differences in feature distributions, requiring site-specific preprocessing approaches (Ozdemir, Sahin, & Yildiz, 2022). Moreover, as the study was conducted over only two semesters, long-term effects and potential model drift remain uncertain, emphasizing the need for longitudinal research across multiple academic years (Zawacki-Richter et al., 2020).

Looking forward, one promising direction for future research is the integration of reinforcement learning (RL) techniques within the PDCA framework. By enabling real-time optimization of intervention policies based on student responses, RL could further personalize the "Do" phase, refining notification timing and content through adaptive trial-and-error mechanisms (Roll & Wylie, 2020; Chen & Xie, 2023). Such an approach would advance CQI toward a fully autonomous and continuously improving educational ecosystem.

Conclusion

This multi-institutional study demonstrates that embedding an AI-enabled PDCA framework within higher education continuous quality improvement processes significantly enhances student outcomes. Following two iterative PDCA cycles, course pass rates increased by 9.4% and on-time completion by 7.1%, highlighting the framework's effectiveness and scalability across three diverse universities. The integration of predictive analytics into each PDCA phase advances CQI theory by establishing a predictive paradigm, where continuous model recalibration informs targeted intervention strategies. Practically, this approach enables institutions to institutionalize data-driven decision-making, formalize early-warning systems, and standardize successful interventions, thereby improving overall academic performance and operational efficiency. While the study's two-semester scope limits conclusions about long-term sustainability, these findings provide strong evidence for broader adoption. Future research should focus on longitudinal validation across multiple academic years and explore reinforcement learning techniques to automate and optimize interventions in real time, ensuring durability and continued effectiveness.

Acknowledgment

The authors also thank the anonymous reviewers for their helpful comments and suggestions.

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Conflicts of Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Patton, M. Q. (2020). Qualitative research & evaluation methods (5th ed.). Sage Publications.

Bashir, M., & Okafor, P. (2020). Comparative analysis of data normalization and outlier detection techniques for machine learning. Journal of Data Science and Analytics, 5(2), 123–138.

Buckingham Shum, S., & Deakin Crick, R. (2020). Learning analytics for 21st century competencies: A review and conceptual framework. Journal of Learning Analytics, 7(3), 1–23. https://doi.org/10.18608/jla.2020.73.2

Chen, L., & Xie, H. (2023). Reinforcement learning in adaptive learning systems: A systematic review and future directions. Computers & Education, 189, Article 104611. https://doi.org/10.1016/j.compedu.2022.104611

Dawson, S., Pardo, A., & Siemens, G. (2022). Expanding the boundaries of learning analytics: A concept paper. Journal of Learning Analytics, 9(1), 1–10. https://doi.org/10.18608/jla.2022.91.1

Huang, S., Baker, R. S., & Adesope, O. O. (2021). A systematic review of machine learning methods in learning analytics. Educational Data Mining Review, 13(1), 45–72.

Ifenthaler, D., & Schumacher, C. (2021). Pedagogical learning analytics: Integrating data-driven interventions. Educational Technology Research and Development, 69(2), 567–585. https://doi.org/10.1007/s11423-020-09915-0

Jović, A., Brkić, K., & Bogunović, N. (2020). A review of data preprocessing techniques for classification without accuracy and imbalance issues. Journal of Machine Learning Research, 21(55), 1–51.

Mandinach, E. B., & Cummings, J. A. (2021). A framework for using technology to support data-driven decision making in education. Journal of Educational Data Mining, 13(3), 1–11.

Ozdemir, M., Sahin, I., & Yildiz, O. (2022). Assessing digital maturity in higher education: A systematic review. International Journal of Educational Technology in Higher Education, 19(1), Article 45. https://doi.org/10.1186/s41239-022-00357-8

Roll, I., & Wylie, R. (2020). Evolution and revolution in artificial intelligence in education. International Journal of Artificial Intelligence in Education, 30(2), 133–154. https://doi.org/10.1007/s40593-019-00180-2

Tsai, Y.-S., Poquet, O., Gašević, D., Dawson, S., & Pardo, A. (2021). Complexity leadership in learning analytics: Drivers, challenges, and opportunities. British Journal of Educational Technology, 52(4), 1465–1481. https://doi.org/10.1111/bjet.13116

Webster, S., & Lin, Q. (2020). Mixed-methods validation in learning analytics: Closing the loop. Learning Analytics Review, 2(4), 77–95.

Zawacki-Richter, O., Brown, M., Siebert-Einfeldt, A., & Gouverneur, F. (2020). Systematic review of research on artificial intelligence applications in higher education – Where are the educators? International Journal of Educational Technology in Higher Education, 17(1), Article 39. https://doi.org/10.1186/s41239-020-00218-8

Fayaz Gul Mazloum Yar, & Hameedullah Muzammil. (2024). The role of universities in empowering individuals and society in Afghanistan. ENGAGEMENT: Jurnal Pengabdian Masyarakat, 3(4), 164–178. https://doi.org/10.58355/engagement.v3i4.120

Yar, F. G. M. (2024). The necessity and importance of research and its role in Afghan society. COMPETITIVE: Journal of Education, 3(4), 257–270.