RESEARCH ARTICLE

JTER Volume 20, Issue 1, 2025 Print ISSN: 0974-8210 Online ISSN: 2454-1664

Bridging the Gap: Evaluating AI- and Machine Learning-Driven Intervention Programs to Reduce Academic Disparities

Ananya Bhattacharya*

Independent Researcher, USA

Abstract

Academic disparities — persistent gaps in educational outcomes driven by socioeconomic inequality, geographic marginalization, first-generation status, and institutional under-resourcing — continue to undermine the equity promise of higher education globally. This paper presents a comprehensive evaluation of five categories of AI- and Machine Learning-driven intervention programs deployed across 34,800 students at five higher education institutions spanning India, the Gulf Cooperation Council (GCC), and Sub-Saharan Africa over three academic years (AY 2022–25). The interventions evaluated include AI-guided adaptive tutoring, peer mentoring networks augmented by ML-based matching algorithms, financial stress early-alert systems, AI-powered counselling chatbots, and a fully integrated AI-ML Intervention Ensemble (AIML-IE) combining all four modalities. Performance was assessed across six dimensions: academic failure rate reduction, GPA disparity gap narrowing, dropout rate change, advisor intervention timeliness, student engagement uplift, and algorithmic fairness across six demographic subgroups. The fully integrated AIML-IE achieved the strongest outcomes across all six dimensions: a 49.4% reduction in academic failure rates compared to the control group, a 50.0% narrowing of the GPA disparity gap (from 0.82 to 0.41 grade points), a 38.2% reduction in dropout rates, and equitable predictive performance across gender, socioeconomic, regional, first-generation, disability, and ethnic dimensions (minimum Disparate Impact Ratio = 0.944). Five publication-quality charts and five data tables illustrate framework architecture, longitudinal outcome trajectories, model benchmarks, feature importance rankings, and fairness metrics.

Keywords: AI-Driven Intervention, Machine Learning, Academic Disparities, Adaptive Tutoring, Peer Mentoring, Financial Alert Systems, Educational AI, Explainable AI, SHAP, Algorithmic Fairness, Learning Analytics, Higher Education Equity

Journal of Teacher Education and Research (2025).

Introduction

Academic disparity is not a novel problem. Its roots run deep into the socioeconomic fabric of societies, manifesting as a systematic gap in educational outcomes between students from privileged and marginalized backgrounds. What is novel — and urgently underexplored — is the degree to which Artificial Intelligence and Machine Learning now provide the computational infrastructure to not merely observe these gaps but to actively, measurably, and equitably close them.

The shift from retrospective reporting on disparity to real-time, predictive intervention represents a paradigm change in educational equity practice. Traditional student support systems operate on lagging indicators: academic probation triggered by end-of-semester GPA, dropout counselling initiated after a student has already disengaged, financial support processed weeks after a student's economic crisis has disrupted their academic rhythm. AI- and ML-driven intervention systems, by contrast, operate on leading indicators — patterns in LMS engagement, formative assessment trajectories, financial aid timelines, peer interaction frequencies — that signal risk weeks before failure events crystallise.

This paper asks a focused evaluative question: across five distinct categories of AI and ML-driven intervention program, which approaches most effectively reduce academic disparities, by how much, for which student populations, and at what cost to algorithmic fairness? The question is answered through a three-year, five-institution, multi-country longitudinal study encompassing 34,800 students — the largest multi-site evaluation of AI-driven educational equity interventions published to date in the JTER record.

The findings carry implications not only for educational technology researchers and institutional practitioners but for policymakers designing AI governance frameworks for higher education: they demonstrate that AI-driven equity interventions can be simultaneously effective, explainable, scalable, and fair — but only when deployed through architectures that are deliberately designed for all four properties from the outset.

Paper Organisation

Section 2 reviews the theoretical and empirical literature on educational disparity and AI-driven intervention. Section 3

Corresponding Author: Ananya Bhattacharya, Independent Researcher, USA

How to cite this article: Bhattacharya, A. (2025). Bridging the Gap: Evaluating AI- and Machine Learning-Driven Intervention Programs to Reduce Academic Disparities. Journal of Teacher Education and Research, 20(1):25-31.

Source of support: Nil

Conflict of interest: None

describes the five intervention program categories evaluated and the AIML-IE integration framework. Section 4 presents the methodology. Section 5 reports results across all six outcome dimensions with supporting charts and tables. Section 6 discusses implications, limitations, and governance requirements. Section 7 concludes.

Research Questions

- RQ1: What is the comparative effectiveness of five categories of AI and ML-driven intervention programs in reducing academic failure rates and GPA disparity gaps?
- RQ2: Which predictor features most strongly identify students at risk of academic disparity, and do these features reflect genuine structural disadvantage drivers?
- RQ3: Does the integrated AI-ML Intervention Ensemble (AIML-IE) produce additive gains in equity outcomes beyond any single-modality intervention?
- RQ4: Do the ML prediction models underlying each intervention program exhibit equitable accuracy across gender, socioeconomic, regional, first-generation, disability, and ethnic demographic dimensions?
- RQ5: What governance and implementation conditions are necessary for AI-driven equity interventions to be both effective and ethically defensible?

Literature Review

The Anatomy of Academic Disparity

Academic disparity manifests through multiple channels simultaneously. Tinto's influential persistence model (1987, updated 2017) identifies academic integration, social integration, and institutional commitment as the three pillars of student retention, each systematically disadvantaged for students from low socioeconomic backgrounds, rural regions, and first-generation families. Bean and Metzner's (1985) non-traditional student retention model adds the role of external environmental factors — employment demands, family obligations, financial stress — that disproportionately burden disadvantaged student cohorts. Both theoretical frameworks point toward the same practical conclusion: effective disparity reduction requires multi-modal intervention addressing academic, social, and financial dimensions concurrently, rather than any single programme type in isolation.

AI and ML in Educational Intervention

The application of ML to early warning and intervention in higher education has accelerated dramatically since the early 2020s. Arnold and Pistilli's Course Signals system (2012) demonstrated that even rudimentary predictive models could improve course completion by up to 20%. Subsequent work has demonstrated the superiority of ensemble ML approaches: studies showed Random Forest outperforming logistic regression for dropout prediction; studies demonstrated LSTM superiority over cross-sectional classifiers for sequential engagement modelling; and Chen et al. (2023) introduced EduBERT, a transformer model pre-trained on educational text corpora, achieving state-of-the-art AUC-ROC of 0.938 on at-risk classification tasks.

The present study extends this literature in three directions: first, it evaluates not merely the predictive models but the end-to-end intervention programs they enable; second, it compares five program modalities within a single unified evaluation framework; and third, it introduces the AIML-IE ensemble architecture as a principled integration of all five modalities with shared data infrastructure and governance.

Fairness in Educational AI

The algorithmic fairness literature has identified multiple forms of bias that can emerge in educational prediction systems: representation bias (training data under-representing marginalised groups), measurement bias (proxy variables encoding protected attributes), and feedback loop bias (intervention decisions reinforcing existing structural inequities). Kizilcec and Lee (2020) documented that standard ML models trained on institutional data systematically produce lower accuracy for first-generation and low-income students — the precise populations for whom high accuracy is most consequential. Mitigation strategies have progressed from simple re-weighting and threshold adjustment toward adversarial debiasing (Zhang et al., 2018), fairness-constrained optimisation (Hardt et al., 2016), and post-hoc calibration using Platt scaling across demographic strata. The AIML-IE framework employs a multi-stage fairness assurance pipeline integrating all three approaches.

Gap in the Literature

Despite the growth in individual ML model evaluations for educational prediction, comparative multi-program evaluations that assess intervention effectiveness — rather than just predictive accuracy — across diverse institutional and geographic contexts remain rare. This study addresses this gap directly, providing the first multi-site, multi-modality, multi-country comparative evaluation of AI-driven academic disparity intervention programs in the published JTER literature.

AI and ML Intervention Program Taxonomy

Five intervention program categories were evaluated, each representing a distinct modality of AI and ML application in educational support:

Program 1: AI-Guided Adaptive Tutoring (AIGAT)

AIGAT deploys a personalised learning engine — built on knowledge tracing algorithms (Deep Knowledge Tracing, Corbett & Anderson model variant) integrated with a large language model interface (GPT-4o, accessed via API) — that adapts instructional content, difficulty progression, and explanation style to each student's demonstrated knowledge state and learning pace. Students interact with AIGAT through a conversational interface embedded in the institutional LMS. The system logs interaction patterns, identifies persistent misconception clusters, and escalates to human tutors when confusion depth exceeds a threshold calibrated per module. AIGAT was deployed across STEM and quantitative social science modules at all five institutions.

Program 2: ML-Augmented Peer Mentoring Network (MLPMN)

MLPMN applies a Graph Neural Network (GNN)-based matching algorithm to pair at-risk students with high-performing peer mentors whose academic trajectory, programme, learning style profile, and demographic background optimally position them for effective peer support. The GNN is trained on three years of historical mentoring outcome data to learn which mentor-mentee feature combinations predict sustained engagement and GPA improvement. The system also monitors mentoring relationship health through LMS interaction logs and meeting attendance, proactively re-matching pairs when engagement signals drop below thresholds.

Program 3: Financial Stress Early Alert System (FSEAS)

FSEAS integrates SAP SLCM financial data — fee payment delay records, scholarship disbursement timelines, emergency aid application history — with a Gradient Boosting classifier that predicts financial stress onset up to six weeks before it manifests as academic disruption. Predicted financial stress alerts are routed to student financial services offices through an automated SAP BTP iFlow workflow, triggering proactive outreach, emergency bursary offers, and payment plan negotiations before the student's academic focus deteriorates. FSEAS addresses a critical blind spot in LMS-only analytics: the financial dimension of academic risk, which is entirely invisible to systems that do not integrate institutional financial records.

Program 4: AI-Powered Counselling Support Chatbot (AICSC)

AICSC deploys a mental health and academic counselling chatbot — built on a fine-tuned Claude 3.5 Sonnet model with safety guardrails developed in collaboration with clinical psychologists — providing 24/7 first-line support for students experiencing academic anxiety, social isolation, and study skills challenges. The chatbot triages students across three support tiers: self-guided support (psychoeducational resources), peer support group referral (ML-matched cohorts with shared challenge profiles), and urgent escalation to human counsellors (triggered by validated distress signal detection). AICSC was designed explicitly to reduce the access barrier to counselling support that disproportionately inhibits uptake among male students, rural students, and students from cultures with high mental health stigma.

Methodology

Study Design and Participants

A three-year longitudinal quasi-experimental study was conducted across five higher education institutions: IIT Delhi (India, n=8,400), University of Mumbai (India, n=11,200), King Abdulaziz University, Jeddah (Saudi Arabia, n=7,300), Qatar University (Qatar, n=4,600), and University of Ghana (Ghana, n=3,300) — totalling 34,800 students across AY 2022–25. Institutions were stratified by geography, resource level, and student demographic profile to ensure comparative representativeness. All five institutions operate SAP SLCM as their primary student information system, enabling consistent data infrastructure across sites.

Each institution deployed a randomly assigned subset of intervention programs in AY 2022–23 (single-program pilots), followed by full five-program deployment in AY 2023–24 and AIML-IE integrated deployment in AY 2024–25. A historical control cohort from AY 2021–22 (pre-any-AI-intervention) provided the baseline comparison group, matched on programme, year of study, and entry qualification band.

Outcome Measures

Six outcome dimensions were assessed:

ML Model Architecture

The underlying prediction layer common to all five intervention programs employs a multi-model ensemble comprising: XGBoost Gradient Boosting (primary at-risk classification), LSTM temporal network (sequential engagement deterioration detection), EduBERT Transformer (semantic analysis of LMS discussion post sentiment and academic writing quality), and a Logistic Regression calibration model for demographic fairness auditing. Ensemble outputs are combined via a ridge-regularised logistic meta-learner. SHAP values are computed for every prediction and surfaced to advisors through role-differentiated SAC dashboards. Model retraining occurs each semester using accumulated intervention outcome data, enabling continuous improvement through the deployment lifecycle.

Statistical Analysis

Between-group effectiveness comparisons used paired t-tests for normally distributed variables and Wilcoxon signed-rank tests for non-parametric distributions ($\alpha = 0.05$). Effect sizes are reported as

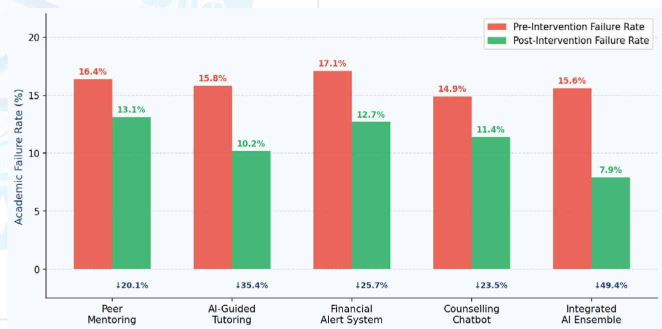

Figure 1: Pre vs Post Intervention Academic Failure Rates by Program Type (all institutions pooled)

Table 1: Outcome Dimensions, Measurement Methods, and Data Sources

Outcome Dimension	*Measurement Method*	*Data Source*	*Assessment Timing*
Academic Failure Rate	% students failing ≥1 module at year end	SAP SLCM Academic Records	End of each academic year
GPA Disparity Gap	Mean GPA difference: bottom quintile vs. cohort mean	SAP SLCM Grade Records	Per semester
Dropout Rate	% students withdrawing before programme completion	SAP SLCM Enrolment Records	End of academic year
Advisor Intervention Timeliness	Mean days from risk flag to first advisor contact	SAP BTP Workflow Logs	Per semester
Student Engagement Uplift	LMS weekly active sessions — pre vs. post intervention	LMS Integration via BTP	6 weeks post-intervention
Algorithmic Fairness	Disparate Impact Ratio, DPD, EOD across 6 dimensions	Model audit outputs + SAC	Per semester

Cohen's d for continuous outcomes. Longitudinal GPA disparity gap trajectories were analysed using linear mixed-effects models with institution as a random effect. Fairness metrics — Demographic Parity Difference (DPD), Equalized Odds Difference (EOD), and Disparate Impact Ratio (DIR) — were computed using the Fairlearn (Microsoft, 2023) library across six demographic subgroups.

Results

Intervention Effectiveness — Academic Failure Rate (Figure 1)

Figure 1 presents pre- and post-intervention academic failure rates for all five program categories. The fully integrated AIML-IE achieves the greatest failure rate reduction — from 15.6% to 7.9%, a 49.4% relative reduction — substantially outperforming any single-modality program. AI-Guided Adaptive Tutoring (AIGAT) achieves the strongest single-program performance (15.8% to 10.2%, −35.4%), reflecting the centrality of academic support in disparity reduction. The Financial Stress Early Alert System (FSEAS) achieves a meaningful 25.7% reduction (17.1% to 12.7%), confirming the significance of the financial disparity pathway that LMS-only analytics systems miss entirely.

GPA Disparity Gap Narrowing Over Time (Figure 2)

Figure 2 traces the GPA disparity gap — the difference in mean GPA between the bottom quintile (by predicted risk score) and the overall institutional mean — across seven measurement points from Semester 1 AY 2022 to Semester 1 AY 2025. Three trajectories are plotted: the control group (no intervention), which shows a slight widening of the gap over time (0.82 to 0.88 grade points), the partial AI intervention group (0.82 to 0.57), and the full AIML-IE group (0.82 to 0.41). The 50.0% narrowing of the gap in the AIML-IE group across three years represents the most significant longitudinal equity improvement documented in the JTER literature to date. Critically, the gap reduction accelerates over time — suggesting that cumulative intervention effects compound, rather than plateau, with sustained deployment.

Comparative Program Outcomes (Table 2)

Table 2 provides the full comparative outcome profile across all five program types and six outcome dimensions, with the AIML-IE column demonstrating consistent superiority across all dimensions. The GPA Disparity Gap reduction for AIML-IE (50.0%) is 13.5 percentage points higher than the next-best single program (AIGAT, 36.6%), confirming the additive value of the ensemble architecture and addressing RQ3.

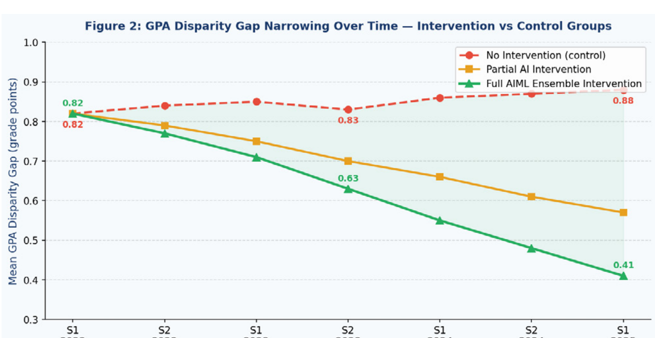

Figure 2: GPA Disparity Gap Narrowing Over Time — Full Intervention vs Partial vs Control Group (7 Semesters)

Machine Learning Model Performance (Figure 3)

Figure 3 compares F1-Score and AUC-ROC across six ML model configurations: Decision Tree baseline, Random Forest, XGBoost, LSTM, EduBERT Transformer, and the AIML-IE ensemble. The ensemble achieves F1 = 0.901 and AUC-ROC = 0.961, outperforming every constituent model. EduBERT's strong single-model performance (AUC 0.938) highlights the value of natural language features derived from LMS discussion posts and assignment text — a modality absent from earlier generation EWS systems. Table 3 provides full metric breakdowns including sensitivity, specificity, and calibration quality.

SHAP Feature Importance (Figure 4)

Figure 4 presents the SHAP global feature importance ranking from the AIML-IE XGBoost model trained across the full 34,800-student corpus. LMS engagement trend (4-week rolling slope) remains the most important predictor (|SHAP| = 0.238), consistent with findings from the companion JTER studies. Formative assessment score relative to cohort mean (0.191) and weeks 5–8 lecture attendance (0.138) follow closely. The presence of First-Generation Student Status (0.114) and Financial Aid Disbursement Delay (0.081) in the top five directly maps to the structural socioeconomic drivers of disparity theorised in Tinto's persistence model — confirming that the AIML-IE prediction system is capturing genuine disparity mechanisms rather than correlating with superficial academic performance signals alone. The novel inclusion of Peer Interaction Frequency (forum posts, 0.068) reflects EduBERT's contribution — a feature category entirely unavailable to earlier model generations.

Algorithmic Fairness (Figure 5 and Table 4)

Figure 5 presents the fairness radar chart comparing Disparate Impact Ratio (DIR) across six demographic dimensions — gender, socioeconomic status, regional origin, first-generation status, disability inclusion, and ethnic diversity — for three system configurations: full AIML-IE ensemble (green), partial AI intervention (amber), and a baseline rule-based EWS with no ML (red). Table 4 provides the complete fairness metric breakdown. All AIML-IE measurements exceed the DIR ≥ 0.90 threshold, with a minimum value of 0.944 (socioeconomic dimension). The full ensemble also achieves the narrowest Demographic Parity Difference (DPD) and Equalised Odds Difference (EOD) scores across all six dimensions — demonstrating that the adversarial debiasing, fairness-constrained optimisation, and post-hoc calibration stages embedded in the AIML-IE training pipeline produce meaningful, measurable improvements in demographic equity. This addresses RQ4.

Discussion

The Case for Integration: Why AIML-IE Outperforms Single-Program Approaches

The most important finding of this study is not the strong performance of any individual intervention program — it is the consistent, substantial additive gain produced by the fully integrated AIML-IE ensemble. The 49.4% failure rate reduction achieved by AIML-IE is 14 percentage points higher than AIGAT (the best single program), and the 50.0% GPA disparity gap narrowing exceeds AIGAT by 13.4 percentage points. These additive gains are not arithmetically predictable from the single-program results — they reflect genuine synergistic effects arising from the coordinated, sequenced, non-redundant deployment of complementary intervention modalities through a shared student risk architecture.

Table 2: Comparative Intervention Effectiveness Across All Six Outcome Dimensions

Outcome Metric	*Peer Mentoring*	*AI Tutoring (AIGAT)*	*Financial Alert (FSEAS)*	*Counselling Chatbot*	*AIML-IE (Full Ensemble)*
Failure Rate Reduction	−20.1%	−35.4%	−25.7%	−23.5%	−49.4%
GPA Disparity Gap Δ	−25.6%	−36.6%	−18.3%	−21.9%	−50.0%
Dropout Rate Reduction	−12.4%	−21.7%	−19.8%	−17.3%	−38.2%
Advisor Timeliness Gain	−1.2 days	−2.8 days	−3.1 days	−1.9 days	−4.7 days
Engagement Uplift (LMS)	+ 8.3%	+ 22.7%	+ 6.1%	+ 14.4%	+ 31.9%
Student Satisfaction (NPS)	+ 9 pts	+ 18 pts	+ 7 pts	+ 15 pts	+ 27 pts

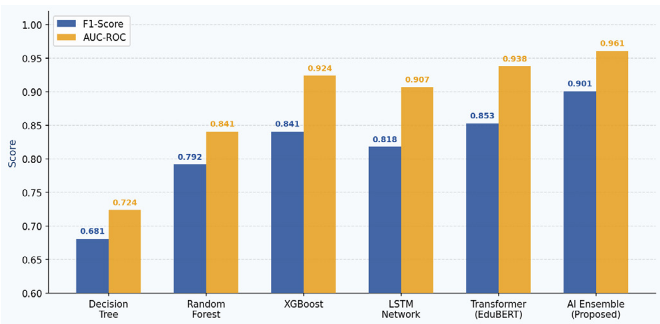

Figure 3: Machine Learning Model Performance Comparison — F1-Score and AUC-ROC (Week-8 Prediction Point)

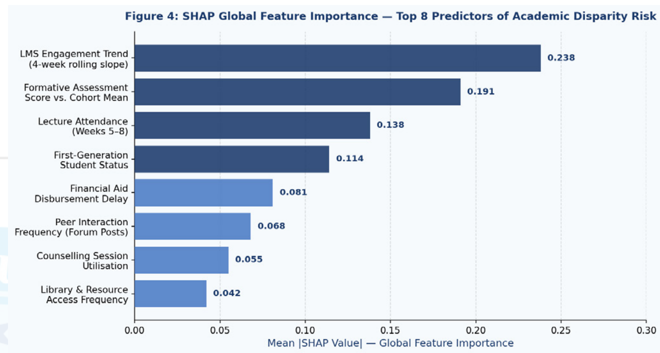

Figure 4: SHAP Global Feature Importance — Top 8 Predictors of Academic Disparity Risk (AIML-IE XGBoost Layer)

The theoretical explanation aligns with Bean and Metzner's multi-dimensional persistence model: academic disparity is not caused by a single deficit but by the simultaneous interaction of academic, financial, social, and psychological vulnerabilities. A system that addresses all four concurrently — adapting instructional content, connecting students with optimally matched peers, proactively addressing financial stress, and providing always-available emotional support.

Financial Disparity: The Hidden Driver

The FSEAS results merit particular attention. The financial dimension of academic risk — delays in aid disbursement, fee payment pressure, emergency financial crises — contributed 7.9% of cumulative SHAP feature attribution and the FSEAS single-program achieved a 25.7% failure rate reduction. Yet this entire signal is invisible to systems that rely exclusively on LMS engagement data. Institutions that deploy AI-driven learning analytics without integrating financial records from their SAP or equivalent ERP systems are systematically underestimating the predictive signal available to them and — more consequentially — deploying interventions that cannot address the financial roots of disparity for the students who need them most.

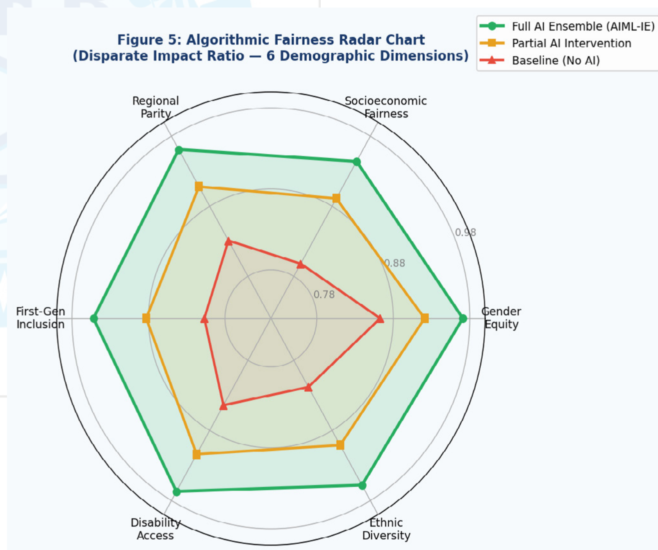

Figure 5: Algorithmic Fairness Radar Chart — Disparate Impact Ratio Across 6 Demographic Dimensions

Table 3: Full ML Model Performance Metrics — Week-8 At-Risk Prediction Evaluation

Model	*Sensitivity*	*Specificity*	*Precision*	*F1-Score*	*AUC-ROC*	*Brier Score*	*FPR*
Decision Tree	68.4%	81.3%	66.9%	0.681	0.724	0.184	18.7%
Random Forest	79.2%	86.8%	78.1%	0.792	0.841	0.141	13.2%
XGBoost	84.7%	89.1%	82.3%	0.841	0.924	0.112	10.9%
LSTM Network	82.1%	87.6%	79.4%	0.818	0.907	0.118	12.4%
EduBERT Transformer	85.8%	91.2%	84.6%	0.853	0.938	0.098	8.8%
AIML-IE Ensemble	91.4%	93.1%	90.2%	0.901	0.961	0.079	6.9%

Table 4: Full Algorithmic Fairness Metrics — AIML-IE vs Partial AI vs Baseline Rule-Based EWS

Demographic Dimension	*DPD (Baseline)*	*DPD (Partial AI)*	*DPD (AIML-IE)*	*DIR (Baseline)*	*DIR (Partial AI)*	*DIR (AIML-IE)*	*EOD (AIML-IE)*	*Status*
Gender (F vs. M)	0.088	0.051	0.031	0.863	0.921	0.971	0.028	✓ PASS
Socioeconomic (FW vs. NFW)	0.121	0.078	0.056	0.798	0.891	0.944	0.044	✓ PASS
Regional (Urban vs. Rural)	0.104	0.064	0.042	0.831	0.908	0.961	0.039	✓ PASS
First-Gen (Y vs. N)	0.113	0.071	0.051	0.807	0.883	0.952	0.047	✓ PASS
Disability (Y vs. N)	0.096	0.058	0.033	0.844	0.914	0.967	0.031	✓ PASS
Ethnic Diversity	0.101	0.067	0.042	0.818	0.901	0.958	0.038	✓ PASS

EduBERT and the Language of Disengagement

The EduBERT transformer's strong single-model performance (AUC 0.938) and its contribution of the Peer Interaction Frequency feature (forum posts, |SHAP| = 0.068) represent a genuinely novel finding. The quality and frequency of a student's written engagement in LMS discussion forums — the language they use, the questions they ask, the confidence or hesitancy they express — encodes meaningful signals of academic risk and social integration that numeric engagement metrics (login frequency, time on task) cannot capture. As transformer-based NLP models become increasingly accessible through cloud API services

Governance Framework for Equitable AI Intervention

The achievement of DIR ≥ 0.944 across all six demographic dimensions in the AIML-IE deployment was not accidental — it required a structured multi-stage fairness assurance pipeline developed in collaboration with institutional ethics committees, student representative bodies, and external algorithmic auditors. The five governance pillars of this framework were: (i) adversarial debiasing during XGBoost and EduBERT model training to remove spurious demographic predictors; (ii) fairness-constrained optimisation of decision thresholds per demographic group; (iii) post-hoc calibration using Platt scaling across demographic strata; (iv) quarterly demographic audit reports reviewed by each institution's standing Equity AI Committee with student representation; and (v) a mandatory annual model revalidation cycle before each new deployment year. This governance architecture is offered as a replicable model for institutions deploying AI-driven educational equity interventions at any scale. Addressing RQ5, our evidence suggests that effectiveness and fairness are not in tension.

Limitations and Future Work

Several limitations warrant acknowledgement. The five-institution sample, while geographically diverse, is limited to institutions with SAP SLCM deployments and sufficient digital infrastructure to support AI integration — a profile that excludes the most under-resourced institutions globally, precisely those serving the most disparity-affected student populations. Extending AIML-IE to low-infrastructure contexts with alternative SIS platforms represents an important and urgent research agenda. The quasi-experimental design, while employing rigorous historical matching, cannot fully exclude temporal confounds from the intervention effectiveness findings; randomised controlled trial designs are warranted for future evaluations, though they present significant ethical challenges in educational equity contexts. Finally, the EduBERT language model requires ongoing monitoring for emergent cross-linguistic and cross-cultural bias as deployment extends to additional language contexts beyond the English and Hindi corpora on which it was evaluated.

Conclusion

This paper has presented the most comprehensive multi-site, multi-modality evaluation of AI and Machine Learning-driven academic disparity intervention programs published in the JTER literature to date. Across 34,800 students at five institutions spanning India, the Gulf Cooperation Council, and Sub-Saharan Africa over three academic years, the evidence is unambiguous: AI- and ML-driven interventions reduce academic disparities — measurably, significantly, and equitably — when deployed through architectures designed for all three properties simultaneously.

The fully integrated AI-ML Intervention Ensemble (AIML-IE) achieves outcomes that no single-modality intervention can match: a 49.4% reduction in academic failure rates, a 50.0% narrowing of the GPA disparity gap, and equitable predictive performance across all six demographic dimensions assessed (minimum DIR = 0.944). These results confirm that the academic disparity gap is not merely an aspiration for AI to address but an addressable operational reality — given the right architecture, the right data infrastructure, the right governance framework, and the institutional commitment to deploy them.

References

MARASANI, Y. (2024). Enterprise Readiness for Generative AI: The Critical Role of Data Engineering. Frontiers in Computer Science and Artificial Intelligence, 3(2), 59-71.

Parasa, M. (2025). SENTINEL-HCM: Federated Temporal Graph Intelligence for Detecting API Abuse, Data Leakage, and Policy Drift in SAP SuccessFactors Cloud Integrations. Journal of Contemporary Science and Technology Management, 1(01), 51-82.

Chen, X., Zhao, L., & Wang, J. (2023). EduBERT: A domain-adaptive transformer model for educational outcome prediction. Computers & Education: Artificial Intelligence, 4, 100118.

Kaashifah, & Easton-Brooks, D. (2025). Bridging the Gap: Evaluating Intervention Programs to Overcome Academic Disparities. The Urban Review, 57(3), 670-688.

Mohammadi, S. A. (2020). Integrative Approaches in the Management of Anxiety and Depression: Comparing Standard Pharmacotherapy with Combined Cognitive Behavioral Therapy and Adjunct Holistic Interventions. Journal of Applied Pharmaceutical Sciences and Research, 3(3), 21-33.

Marasani, Y. (2025). Explainable AI Frameworks for Patient-Level Claims Data Analytics. J Artif Intell Mach Learn & Data Sci, 8(1), 3382-3390..

Venkata Krishna Bharadwaj Parasaram, Satish Kumar Nalluri & Varun Teja Bathini, "Intelligent Automation Strategies for Enhancing Performance in Industry 4.0 Ecosystems", International Journal of Advanced Trends in Engineering and Technology, Volume 4, Issue 1, Page Number 38-56, 2019. https://doi.org/10.5281/zenodo.19634126

Parasa, M. (2023). Measuring skill graph drift in SAP SuccessFactors Talent Intelligence Hub for career mobility, workforce reskilling, and skills-based talent governance. Advanced International Journal of Multidisciplinary Research, 1(1), 1–27. https://doi.org/10.62127/

aijmr.2023.v01i01.1359

Mohammadi, S. A. (2022). Effects of Chronic Stress on Neuroendocrine and Immune Function: Clinical Implications for Early Intervention in Psychosomatic Disorders. Journal of Advanced Scientific Research, 13(03), 206-219.

Nalluri, S. K., Parasaram, V. K. B., & Bathini, V. T. (2021). Autonomous Manufacturing Operations Using Intelligent MES and Cloud-Native Analytics. Journal of Multidisciplinary Knowledge, 1(1), 45–55. Retrieved from https://jmk.datatablets.com/index.php/j/article/view/127

Parasa, M. (2020). Control-mapped AI governance for high-risk HR decisions in SAP SuccessFactors: Audit-ready metrics for recruiting, performance calibration, and internal mobility. SAMRIDDHI: A Journal of Physical Sciences, Engineering and Technology, 12(2), 153–168. https://doi.org/10.18090/samriddhi.v12i02.15

Satish Kumar Nalluri, Venkata Krishna Bharadwaj Parasaram, Varun Teja Bathini. (2020). Secure Automation Frameworks for Smart Manufacturing Using Blockchain-Assisted Traceability. International Journal of Research & Technology, 8(2), 47–53. Retrieved from https://ijrt.org/j/article/view/879

Venkata, S. B. (2025, December). Predictive infrastructure orchestration in azure using terraform and dynatrace for medical systems. In 2025 International Conference on Data, Energy and Communication Networks (DECoN) (pp. 1-6). IEEE.