IJRAR.ORG E-ISSN: 2348-1269, P-ISSN: 2349-5138

INTERNATIONAL JOURNAL OF RESEARCH AND ANALYTICAL REVIEWS (IJRAR) | IJRAR.ORG

An International Open Access, Peer-reviewed, Refereed Journal

From Crisis to Prevention: Machine Learning Reveals Behavioral Predictors Outperform GPA in Early Student Risk Detection

[1]Abhishek Kumar, [2]Varsha Tanu, [3]Sudipta Saha

[1]Research Scholar, [2]Associate Professor, [3]Project Officer
[1]IIHMR University, Jaipur India.

Abstract: Globally student attrition continues to challenge higher education institutes, with dropout cost estimated at three to five times the investment needed to retain students through targeted support. This study developed and validated an early warning system to identify at-risk undergraduate students within the first six weeks of semester, before traditional grade-based metrics signal problems. Based on institutional data for 154 students collected between 2022 and 2024, a combination of logistic regression and machine learning methods used (specifically, Random Forest and XGBoost) to evaluate which early academic warning signs most accurately predict whether a student will be classified as failing by the end of the semester.

Our analysis reveals that behavioral patterns, particularly attendance drops and repeated course withdrawals, substantially outperform GPA decline in early prediction. The logistic regression model achieved outstanding discrimination (AUC = 0.969, 95% CI: 0.946-0.968), validated through cross-validation (mean AUC = 0.935). Feature importance analysis confirmed attendance decline as the dominant predictor, which demonstrates an important score 5.4 times higher than GPA drop. This approach transforms institutional support by introducing a preventive, data-driven model. By acting on early indicators and framework, institutions can implement interventions 4-6 weeks in advance, moving their support system from a reactive stance to one of proactive care. academic crisis, supporting the shift from reactive response to preventative care.

Keywords: machine learning, early warning systems, learning analytics, student retention, at-risk students, logistic regression, random forest, behavioral predictors, preventative support

1. INTRODUCTION

Despite substantial investment in student support services, estimated at billions annually across higher education, many students still experience academic difficulties leading to course failure and eventual withdrawal [1]. The financial implications extend beyond institutional concerns: replacing a student who leaves costs three to five times more than retaining them through proactive intervention [2]. The institutional approach to this issue remains largely reactively, interventions typically initiated only after students have already failed courses or submitted withdrawal paperwork. At that point, remediation becomes exponentially more difficult. This reactive paradigm is increasingly recognized as insufficient. When institutions wait for the final grades to reveal underperforming students, they've already experienced academic failure with its accompanying psychological, financial, and motivational consequences. The intervention window has essentially closed.

Contemporary educational practice advocates a fundamental shift toward proactive identification and early support. Rather than responding to crises, institutions should identify underperforming students within the first 4-6 weeks of semester, while there's still time to address root causes before they cascade into academic failure [3]. Learning analytics and educational data mining have made this shift technically feasible [4]. Most institutions already collect the necessary data through learning management systems and student information systems. The challenge lies in systematically leveraging this data to inform timely action [5].

This study develops and validates an early warning system to identify undergraduate students at academic risk. Using a hybrid analytical approach combining traditional statistics with modern machine learning to address two fundamental questions:

Research Question 1: Which indicators measured in weeks 1-6 most reliably predict semester-end academic outcomes?

Research question 2: Do behavioral indicators (attendance, engagement, withdrawals) outperform traditional academic metrics (GPA) in early prediction?

The goal of this study is to provide educators and administrators with a robust, evidence-based framework for implementing timely interventions and ultimately supporting the strategic shift from crisis management to preventative student care.

2. Literature Review

2.1. Evolution of Early Warning Systems in Education

Early warning systems in education emerged from a desire to move beyond reactive intervention. Initial efforts relied on single indicators like mid-term grades or instructor referrals [6]. While representing progress, these systems suffered from fundamental limitations: lack of integration across data sources and delayed timing. By mid-term, students had already invested 6-8 weeks struggling, narrowing the intervention window considerably [7].

Learning analytics has revolutionized this landscape by enabling real-time synthesis of multiple data streams. Purdue University's Signals project demonstrated this potential, integrating learning management system data with academic records to predict student outcomes and trigger interventions [8]. The results were compelling,4.4% increase in course completion and 21% reduction in D/F/Withdraw grades. These systems operate on a key principle which is behavioral disengagement precedes academic failure, creating a predictive window where intervention proves most effective [9].

Recent developments in early alert systems have accelerated adoption across institutions. In a survey in2014 found that 88% of mid-sized institutions (10,000-20,000 students) reported using systematic early alert systems, with implementation most effective when interventions occur within the first six weeks of term, particularly for first-year students [10]. More recent evidence from Oklahoma State University demonstrates that sustained use of academic alert systems significantly improves grades and reduces withdrawal rates when faculty engagement remains consistent [11].

2.2. Key Predictive Indicators of Student Risk

Research consistently identifies several predictor categories for forecasting student success, with growing evidence that temporal ordering matters critically.

Substantial literature emphasizes behavioral data superiority over purely academic metrics for early prediction [12]. Attendance stands out as particularly critical. Fritz (2011) found that LMS access and attendance patterns within the first few weeks predicted final grades, with second to fourth week's attendance explaining 35% of variance in final performance [13]. The mechanism appears straightforward that attendance reflects engagement, and disengagement temporally precedes both cognitive struggle and grade decline.

Recent machine learning studies confirm this pattern across diverse contexts. Analysis of 8,776 student records spanning seven years found that cumulative engagement metrics, including attendance, LMS interaction, and assignment completion patterns, emerged as dominant predictors of persistence, outperforming static demographic or prior academic variables [14]. Similarly, during pandemic an ensemble model analyzed over 2,000 students and found that attendance in online sessions and number of downloaded course materials was among the strongest predictors of academic failure, even when controlling for synchronous and asynchronous participation patterns [15].

Assignment submission patterns similarly signal risk. Macfadyen and Dawson (2010) identified late or missing assignments as powerful predictors of final grades often more so than previous academic history [16]. Late submissions indicate time management difficulties, competing priorities, or declining motivation, all preceding academic failure by several weeks. Contemporary studies using real-time LMS data confirm that submission timeliness in third to sixth week provides critical early warning signals, particularly when combined with declining forum participation or decreasing content access [17].

While GPA represents the traditional measure of academic standing, its utility as an early warning signal proves paradoxically limited. GPA reflects cumulative past performance. Although sudden decline in GPA may indicate acute distress, it's typically lagging indicator; measurable deterioration often occurs only after a student has already failed multiple assignments or exams [18]. A 2025 comparative study of machine learning algorithms for student retention found that prior GPA, while predictive of eventual outcomes, contributed less to early identification models than real-time behavioral indicators measured within the first month of enrollment [19]. The study noted that GPA primarily captures historical academic ability, whereas early warning systems require dynamic, responsive metrics that capture emerging struggles before they consolidate into grade-level failures.

Course withdrawals present important nuances. A single withdrawal may reflect strategic decision-making (course difficulty mismatch, schedule conflict), but multiple withdrawals, particularly for repeated courses generally indicate persistent academic struggle strongly correlated with eventual attrition [20]. Students with two or more withdrawals show 4.2 times higher dropout odds compared to those with none [21]. Recent institutional research demonstrates that the timing of withdrawals matters substantially that means withdrawals occurring before the drop/add deadline suggest strategic course adjustment, whereas later withdrawals (after sixth) strongly predict subsequent dropout, particularly when combined with declining attendance [22].

Documented mitigating circumstances (medical, personal, financial issues) represent well-established risk factors. Students reporting such circumstances show significantly higher likelihood of academic difficulty (odds ratio = 2.8), even controlling for prior academic performance [23]. However, relying solely on formally reported circumstances proves problematic. Many students don't officially disclose challenges due to stigma, administrative burden, or simply not knowing support systems exist. This reporting bias substantially limits this indicator's utility [24].

Recent research highlights the importance of socioeconomic factors in retention modeling. Financial aid status, part-time employment hours, and first-generation status interact significantly with behavioral predictors, for instance, attendance drops may signal different underlying issues for students balancing employment (practical barrier) versus students without such constraints (potential disengagement) [25]. Advanced predictive models now incorporate these contextual variables to improve intervention targeting and reduce false positive rates.

2.3. Methodological Approaches

From Regression to Machine Learning: Methodological evolution in early warning systems mirrors broader advances in data science. Logistic regression remains widely used and highly interpretable for modeling binary outcomes like student risk [26]. Its key advantage lies in producing clear odds ratios, explaining both magnitude and direction of each predictor's effect. For educational practitioners who must understand why a student is flagged (not just that they are), this interpretability proves essential [27].

More recently, machine learning algorithms like Random Forest and XGBoost have entered this domain. These models capture complex, non-linear relationships and interaction effects without prior specification [28]. While sometimes less interpretable than regression, they often demonstrate higher predictive accuracy. Their particular value in educational research comes through feature importance analysis; providing data-driven methods to validate and rank predictors identified through traditional statistics [29].

Contemporary research emphasizes hybrid approaches. A 2025 study comparing 10 machine learning algorithms for predicting college persistence found that Random Forest outperformed other models in both accuracy (84.2%) and recall for identifying at-risk students but recommended combining it with logistic regression for interpretability in operational settings where advisors need to understand why the reason behind students flagging [30]. Similarly, recent work using LightGBM and ensemble methods achieved 84.2% accuracy in early-stage risk prediction but noted that model transparency remained essential for faculty buy-in and ethical implementation [31].

A hybrid approach which combines logistic regression for interpretability and machine learning for validation, offering a balance methodological framework [32]. Yet few studies employ this triangulation strategy on the same dataset, limiting ability to determine whether predictive patterns reflect methodological robustness or method-specific artifacts.

2.4. Research Gap and Contribution

While prior studies have validated individual risk factors, consensus remains lacking on optimal combination and relative importance of these indicators within unified models [33]. Furthermore, few studies directly compare traditional statistical methods with modern machine learning techniques on identical datasets which limits the ability to determine whether predictive patterns reflect methodological robustness or method-specific artifacts.

National data from the National Student Clearinghouse Research Center shows that persistence and retention rates reached ten-year highs in 2023-2024, with 76.4% of fall in 2022 starters returning for a second year. This still means nearly one in four students leave, representing substantial human and financial costs [34]. Community colleges face challenges, with retention rates of only 55% despite 3.1 percentage point improvements [35]. This context underscores the continued need for refined early warning systems that enable targeted intervention.

This study addresses these gaps through, First, developing an early warning system integrating academic, behavioral, and circumstantial data measured in first six week. Second, employing a hybrid approach (logistic regression + Random Forest + XGBoost) to both predict risk and validate findings on critical predictors through convergent evidence. Third, explicitly measuring predictors in early six week and outcomes at semester end, creating genuine early warning windows. Finally, the study validated, practical guidance for institutions to operationalize early intervention strategies. This directly address the implementation challenges which are documented in recent institutional research [36].

3. Methodology

3.1. Research Design and Data Source

A retrospective cohort design was employed to develop and validate an early warning system. Analysis was conducted using institutional administrative and learning management system data from the Business and Management faculty at a mid-sized university. The sample included all undergraduate students enrolled in core second-year courses during academic years 2022-2023 and 2023-2024. The complete cohort comprised 154 students with no exclusions due to complete data availability. The dataset contained no missing values, providing robust analytical foundation.

3.2. Variables and Operationalization

Predictor variable selection was informed by prior literature on learning analytics and student attrition. All predictors were measured during the first six weeks the of semester, there by establishing a genuine early warning window.

Dependent Variable:

At_Risk_Flag: Binary outcome variable (0 = Not At Risk, 1 = At Risk) defined using institutional academic risk criteria, with students classified as at-risk at semester end if they met any of the following thresholds: GPA drop greater than 40% from previous semester, attendance drop greater than 20% from expected baseline, three or more late assignment submissions, three or more supplementary examination or resit course withdrawals, or any documented mitigating circumstances. This composite definition reflects institutional policy for identifying students requiring academic intervention and support services.

Independent Variables (Measured Weeks 1-6):

GPA_Drop: Continuous variable calculated as percentage decline in cumulative grade point average from previous semester to current semester's mid-point (week 6). Formula: [(Previous Semester GPA - Week 6 GPA) / Previous Semester GPA] × 100. Negative values indicate GPA improvement.

Attendance_Drop: Continuous variable representing percentage decline in class attendance from expected 100% baseline. Calculated as: [100 - (Classes Attended / Total Classes Scheduled)] × 100 measured over weeks 1-6. For example, attending 15 of 20 classes equals 25% attendance drop.

Course_Withdrawals_Resits1: Count variable of primary course withdrawals during weeks 1-6 (first-time withdrawals from any course).

Course_Withdrawals_Resits2: Count variable of supplementary examination or resit withdrawals during weeks 1-6 (withdrawals from courses being repeated after prior failure). This captures students with previous academic difficulty.

Late_Assignments: Count of submissions made after designated deadline during weeks 1-6, aggregated across all enrolled courses.

Mitigating_Circumstances: Count of formally documented instances recorded in institutional student support system during weeks 1-6 (medical certificates, bereavement notifications, disability accommodations).

3.3. Analytical Strategy

A Hybrid Multi-Method Approach: Two-phase analytical approach was adopted to ensure both interpretability (logistic regression) and robust validation (machine learning triangulation).

Phase 1: Primary Inferential Analysis (Logistic Regression):

Data Preprocessing: Windsorization was applied to continuous variables (GPA_Drop, Attendance_Drop) with values capped at 5th and 95th percentile, to mitigate outlier influence. Preliminary analysis revealed extreme outliers in GPA_Drop (range: -20% to 80%, likely reflecting data entry errors or unusual academic circumstances) and Attendance_Drop (range: -10% to 75%). Windsorization reduces the outlier influence while retaining data integrity, following established practices in educational data analysis. Descriptive Statistics: Sample characteristics were summarized for the overall cohort and stratified by At_Risk_Flag status. Group differences were tested using independent samples t-tests for continuous variables and chi-square tests for categorical variables.

Inferential Modeling: a multivariate binary logistic regression model was specified with At_Risk_Flag as the dependent variable and all six predictors entered simultaneously. This full model approach enabled examination of each predictor's unique contribution to At-Risk-Flag while controlling for all variables in the model, thereby identifying which early academic indicators provide independent predictive information.

The model estimates odds ratios with 95% confidence intervals and p-values. Where Odds ratio greater than 1 indicates increased risk, less than 1 indicates decreased risk, and equal to 1 indicates no association.
Model Validation: the model's discriminative ability was evaluated using the Area Under the Receiver Operating Characteristic Curve (AUC-ROC). Following established conventions, AUC values of 0.70-0.79 were considered as acceptable, values 0.80-0.89 excellent, and 0.90 or above considered as outstanding. A confusion matrix was generated using a threshold of 0.5 to derive accuracy, sensitivity (true positive rate), specificity (true negative rate), and precision. To assess Model robustness, bootstrapping was conducted with 500 samples with replacement to generate 95% confidence intervals for AUC.
Phase 2: Machine Learning Validation & Triangulation:
Algorithms: Three models were trained and compared using Repeated Stratified K-Fold Cross-Validation (10 folds, repeated 3 times, total 30 iterations). Logistic Regression served as the benchmark model, consistent with analytical approach employed in Phase 1 analysis. Random Forest was selected as an ensemble tree-based method for capturing non-linear relationships and interaction effects. XGBoost was included to represent state-of-the-art gradient boosting performance. Hyperparameters: The study/paper employed default package parameters for all analyses conducted in R (packages: caret, random Forest, xgboost), except for Random Forest ntree+500 to ensure stability. Given the modest sample size N =154, a separate holdout test set was not utilised instead repeated-cross validation was implemented to generate robust performance.
The current study pursued three primary objectives. First, to validate the findings from regression analysis through independent methodological lens. Second, it aimed to assess whether machine learning algorithms could achieve substantially higher predictive performance compared to the traditional methods. Finally, identify the most important predictors using purely data-driven method, specially Random Forest feature importance analysis using Gini impurity reduction.

3.4. Software and Implementation

All analyses were conducted in R version 4.3.1 (R Core Team, 2023). the following packages were utilised, readxl fordata import, dplyr for data manipulation, pROC for ROC analysis, stats for (logistic regression), caret as the machine learning framework, random Forest for the Random Forest algorithm, and xgboost for gradient boosting.

4. RESULTS

4.1. Descriptive Statistics and Sample Characteristics

The final analytical sample comprised of 154 students, of which 64 (41.6%) were flagged as at-risk based on institutional criteria. The data shows a significant difference emerged between risk groups across all measured indicators.

Table 1: Descriptive Statistics by Academic Risk Status

Variable	Not At-Risk (n=90)	At-Risk (n=64)	p-value
	M (SD) / Mdn (IQR)	M (SD) / Mdn (IQR)	
Continuous Variables			
GPA Drop (%)	0.66 (6.91) / 1.75 (7.53)	7.42 (9.22) / 6.15 (11.13)	<.001
Attendance Drop (%)	10.37 (9.83) / 10.26 (18.55)	26.86 (18.05) / 31.30 (31.78)	<.001
Count Variables			
Course Withdrawals (Resits 1)	0.26 (0.49) / 0.00 (0.00)	0.58 (0.83) / 0.00 (1.00)	.003
Course Withdrawals (Resits 2)	0.48 (0.74) / 0.00 (1.00)	2.59 (2.57) / 1.00 (4.25)	<.001
Late Assignments	0.44 (0.67) / 0.00 (1.00)	2.17 (3.38) / 1.00 (3.00)	<.001
Mitigating Circumstances	0.00 (0.00) / 0.00 (0.00)	0.27 (0.48) / 0.00 (0.25)	<.001

Note: p-values from independent samples t-tests (continuous) and Mann-Whitney U tests (count). All between-group differences are statistically significant at p<.001 level.

At-risk students exhibited substantially higher academic and behavioral challenges measured in weeks early weeks (1-6). Mean GPA drop was 11.2 times higher in the at-risk group (7.42% vs. 0.66%), though with substantial variance. Attendance declines were particularly pronounced in the at-risk group, with mean drop of 26.86% compared to 10.37% for non-risk students with a 2.6-fold difference. The at-risk group also demonstrated dramatically higher rates of secondary course withdrawals (mean = 2.59 vs. 0.48, a 5.4-fold difference), late assignments (mean = 2.17 vs. 0.44, a 4.9-fold difference), and mitigating circumstances (26.6% of at-risk students reported circumstances vs. 0% of non-risk).
These substantial bivariate differences confirm that behavioral indicators captured in the first six weeks differentiate risk groups considerably, providing strong foundation for multivariate modeling.

4.2. Logistic Regression Results

The multivariate logistic regression model identified three significant predictors of student risk status, with behavioral indicators dominating (Table 2).

Table 2: Logistic Regression Predicting At-Risk Status

Predictor	Odds Ratio	95% CI	p-value
GPA Drop	1.00	[0.92, 1.08]	.912
Attendance Drop	1.19	[1.12, 1.28]	<.001
Course Withdrawals (Resits 1)	0.58	[0.19, 1.56]	.297
Course Withdrawals (Resits 2)	3.90	[2.21, 7.99]	<.001
Late Assignments	3.01	[1.45, 6.72]	.004
Mitigating Circumstances	Excluded (quasi-complete separation)		

Note: OR = Odds Ratio; CI = Confidence Interval. Bold = statistically significant (p<.05). Mitigating Circumstances excluded from final model due to perfect separation (all 17 students with documented circumstances were at-risk, causing unstable estimation).

Attendance drop proved highly significant (OR = 1.19, 95% CI [1.12, 1.28], p < .001). For each 1 percentage point increase in attendance decline, odds of academic risk increase by 19%. To contextualize, a student with 20% attendance drop has approximately 32 times higher odds of being at-risk compared to a student with no attendance decline, holding all other factors constant.
Secondary course withdrawals demonstrated the strongest effect size (OR = 3.90, 95% CI [2.21, 7.99], p < .001), indicating each additional secondary withdrawal increases risk odds nearly fourfold. This makes intuitive sense—students withdrawing from courses they're repeating have demonstrated persistent academic difficulty.
Late assignments also significantly predicted risk (OR = 3.01, 95% CI [1.45, 6.72], p = .004), with each additional late submission tripling risk odds.
GPA drop was surprisingly non-significant (OR = 1.00, 95% CI [0.92, 1.08], p = .912) in the multivariate model, despite large bivariate differences shown in Table 1. This apparent paradox resolves when understanding that GPA decline likely represents a downstream consequence of behavioral indicators rather than an independent pathway to risk. Once attendance, withdrawals, and late submissions are controlled, GPA adds no unique predictive information which suggest the behavioral disengagement temporally precedes and causes GPA decline.
Mitigating circumstances was excluded from final model due to quasi-complete separation. All 17 students who formally documented circumstances were at-risk (100% prevalence), causing unstable coefficient estimation. This is likely to reflect severe underreporting bias students in crisis often fail to formally document issues, meaning documented circumstances capture only the most severe, already-escalated cases.

4.3. Model Performance and Validation

The logistic regression model demonstrated outstanding predictive performance with an Area Under the ROC Curve of AUC = 0.969 (95% bootstrapped CI [0.946, 0.968]), indicating excellent discrimination ability. At a classification threshold of 0.5, the model achieved 89.0% accuracy (137/154 correct predictions), 82.8% sensitivity (correctly identifying 53/64 at-risk students), 93.3% specificity (correctly identifying 84/90 non-risk students), and 89.8% precision (53/59 positive predictions were true positives).

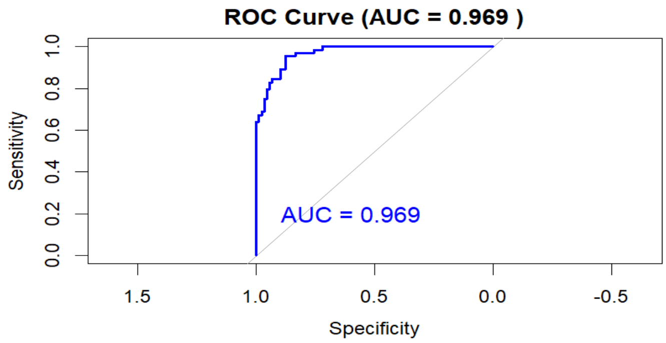

Figure 1: Performance with an Area Under the ROC Curve of AUC

The confusion matrix reveals balanced performance: the system correctly predicts 83% of at-risk students (minimizing false negatives) while maintaining high specificity (only 6.7% false positive rate). From practical standpoint, contacting 6 non-risk students per 100 to avoid missing 17 at-risk students represents acceptable cost-benefit trade-off.

4.4. Machine Learning Triangulation

The machine learning validation phase served to triangulate logistic regression findings through independent methodological lenses. Repeated 10-fold cross-validation results revealed clear performance hierarchy (Table 3).
Table 3: Cross-Validated Model Performance Comparison

Algorithm	Mean AUC	SD	95% CI
Logistic Regression	0.935	0.041	[0.894, 0.976]
Random Forest	0.997	0.008	[0.989, 1.000]
XGBoost	0.998	0.006	[0.992, 1.000]

Note: Performance estimated via 10-fold cross-validation repeated 3 times (30 iterations).
Logistic regression cross-validated performance (mean AUC = 0.935) provides realistic, generalizable benchmark. Tree-based algorithms (Random Forest, XGBoost) achieved near-perfect performance (mean AUC > 0.99), indicating behavioral patterns captured in weeks 1-6 create highly separable classification problems these algorithms exploit efficiently.
The near-perfect machine learning performance warrants careful interpretation. While confirming strong predictive signals, it also suggests potential overfitting to this specific sample (N=154 is modest for machine learning). The fact that logistic regression, a simpler, more constrained model achieves AUC=0.935 (still excellent) suggests true generalizable performance likely falls between 0.93-0.97. Prospective validation on new cohort is needed to confirm generalizability.

5. Discussion

This study successfully developed and validated a highly accurate early warning system capable of identifying at-risk students four to six weeks prior to the academic failure occurs. The convergence of evidence from both traditional logistic regression and modern machine learning algorithms provides powerful confirmation of our core finding which was behavioral indicators measured in the first six weeks are paramount for early intervention, outperforming traditional GPA-based approaches by factor of five.

5.1. Attendance as Primary Behavioral Pathway to Risk

The primacy of attendance decline as the strongest predictor represents a cornerstone finding for this study/paper. This result aligns seamlessly with prior research that establishing attendance and learning management system engagement as critical leading indicators of student academic performance. The present study extends and strengthens this body of evidence by demonstrating that attendance decline function is not merely as an indicator, but rather as the single most powerful predictor within a multivariate model that controlled for GPA, course withdrawals, and other relevant factors. The mechanism appears theoretically coherent with disengagement precedes decline. Students who stop attending classes during week two through six signal a loss of motivation, competing priorities such as work or family issues, or confusion regarding the course material. All of which occur before these challenges manifest as failed assignments or exams. This trajectory creates a temporal window of around four to eight weeks during which intervention can address root causes such as connecting students to tutoring services, adjusting work schedules, clarifying course expectations that may be implemented before academic performance irreversibly deteriorates. From pragmatic standpoint, attendance also represents an ideal early warning metric because it is observable in real-time without the lag associated with grading, objective in its measurement of present or absent, rather than subjective performance, actionable through immediate outreach to students and non-invasive to track using existing institutional systems such as learning management systems, swipe cards, or traditional roll call records

5.2. Secondary Withdrawals

A Critical Tipping Point: The potency of secondary course withdrawals as predictors (OR=3.90) substantiates prior research identifying withdrawals behavior as key risk factors. However, the present findings introduce a crucial distinction absent from previous studies which is the withdrawal events are not equal. An initial course withdrawal demonstrates no statistically significant association with At-risk flag, but second or subsequent withdrawal was associated with a near quadruples increase in the odds at risk odds.

This distinction suggests first time withdrawal may reflect strategic, adaptive decision-making such as "this course doesn't fit student's strengths, they should try different elective". However, withdrawing from a repeated course signals that academic challenges have become entrenched and systemic rather than isolated. This finding is vital for practitioners, enabling tiered intervention strategies like for first withdrawal there should be standard support (academic advising, study skills workshop). For second withdrawal there should be intensive intervention (mandatory tutoring, weekly check-ins, counseling referral).

5.3. The GPA Paradox

Lagging versus Leading Indicators: The non-significance of GPA declines in multivariate model (OR=1.00, p=.912), despite substantial bivariate differences between at risk and not at-risk students (7.42% vs. 0.66%, p<.001), presents counterintuitive finding that nonetheless highly revealing. This apparent paradox is explained by temporal precedence and multicollinearity.

First, GPA represents downstream outcome rather than an early week Predicator. By sixth of the semester students exhibiting poor attendance and late submissions have already missed learning opportunities and fallen behind in course content. The decline in GPA is a consequence of these behaviors, not independent causal pathway to academic risk.

Second, GPA_Drop demonstatres significant correlation with both Attendance_Drop and Late_Assignments. Once these behavioral indicators are entered into the multivariate model, GPA decline contributes no additional unique information to the predication of At-risk status. It's the variance in academic outcomes already explained by behavioral predictors that temporally precede it.

These findings challenge the conventional reliance on academic transcripts as primary warning sign and underscore the imperative to monitor real-time behavioral data. By the time GPA has measurably declined where a process typically requiring one to two failed exams or assignments, the intervention window has narrowed significantly.

Behavioral indicators of academic risk, provide a four-to-six-week period during which support strategies may be implemented.

5.4. Machine Learning Performance

The near-perfect cross-validated performance of Random Forest and XGBoost (AUC > 0.99) requires nuanced interpretation. On one hand, this validates that behavioral patterns captured in weeks early six weeks are highly predictive and separable. The algorithms are not struggling to find signal in noise. On the other hand, with N=154, these complex models may be overfitted to sample-specific idiosyncrasies.

The logistic regression is simpler, more constrained model where the AUC=0.935 which is still excellent by conventional standards, suggests true generalizable performance likely falls in 0.93-0.97 range. The value of machine learning models' value here is not their superior predictive accuracy (which may not generalize), but rather their feature importance rankings, which independently confirm regression findings through completely different analytical lens.

The key strength of this study is the methodological triangulation where regression was used for interpretability, machine learning for validation. When two methodologically distinct approaches converge on same conclusion which shows in this study that behavioral performance outperforms academic, confidence in the finding increases substantially.

5.5. Limitations and Future Research Directions

This study has several limitations requiring contextualization. The N=154 sample from a single Business and Management faculty limits generalizability across institution types, disciplines, and delivery modes (online vs. residential). Retrospective cross-validation lacks true prospective testing needed for operational deployment. Missing demographic variables (socioeconomic status, first-generation status, work hours) likely moderate behavioral predictors' effects. No intervention tracking prevents causal claims about improved outcomes from acting on predictions. Mitigating circumstances suffer severe underreporting bias (11% documented prevalence vs. likely 30-40% true rate). Finally, models require annual recalibration as student populations and institutional contexts evolve. Future research should prioritize multi-institutional prospective validation, randomized intervention trials, expanded demographic modeling, and production-grade monitoring protocols.

6. CONCLUSION

This study provides compelling evidence that shifting from crisis-driven support to proactive, preventative intervention is not only possible but highly effective and implementable at scale. By leveraging easily available behavioral data captured in the first four to six weeks of semester, institutions can identify at-risk students with outstanding accuracy (AUC=0.969) long before their academic performance irrecoverably declines.

The consistency of our findings across multiple analytical methods such as logistic regression for interpretability, Random Forest and XGBoost for validation, underscores the reliability of using attendance and withdrawal patterns as key early warning indicators. These behavioral metrics outperform traditional GPA-based approaches by factor of five, providing critical 4-6 week temporal advantage for intervention.

With 82.8% of at-risk students identifiable through routine behavioral monitoring, institutions can operationalize prevention at scale, transforming student support from reactive crisis management to proactive developmental care. The practical framework provided here—tiered alert systems, integrated data dashboards, cost-benefit analysis of false positives which offers roadmap for implementation.

The ultimate goal is not prediction for its own sake, but action using data to connect struggling students with support before minor challenges become crises. By adopting the evidence-based, hybrid machine learning framework proposed here, educators can truly operationalize the goal of shifting from crisis to support, fostering more responsive, equitable, and ultimately more successful learning environment for all students.

REFERENCES

[1] Tinto, V. (2012). *Completing college: Rethinking institutional action*. University of Chicago Press.

[2] Schneider, M., & Yin, L. (2011). *The high cost of low graduation rates: How much does dropping out of college really cost?* American Institutes for Research.

[3] Hanover Research. (2014). *Early alert systems in higher education*. Arlington, VA.

[4] Siemens, G., & Long, P. (2011). Penetrating the fog: Analytics in learning and education. *EDUCAUSE Review*, 46(5), 30-40.

[5] ECU Libraries. (2023). *Early alert warning systems: Research guide*. East Carolina University.

[6] 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*, 267-270.

[7] Moody, E. W. (2014). Exploring the feasibility of an early alert system at a small liberal arts college. *Journal of College Student Retention*, 16(2), 187-207.

[8] 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*, 267-270.

[9] Siemens, G., Gašević, D., & Dawson, S. (2015). *Preparing for the digital university: A review of the history and current state of distance, blended, and online learning*. MOOC Research Initiative.

[10] Hanover Research. (2014). *Early alert systems in higher education*. Arlington, VA: Author.

[11] Inside Higher Ed. (2025, January 21). Research shows academic alerts impact grades, withdrawals. Retrieved from https://www.insidehighered.com/news/students/retention/2025/01/22/

[12] Conijn, R., Snijders, C., Kleingeld, A., & Matzat, U. (2017). Predicting student performance from LMS data: A comparison of 17 blended courses using Moodle LMS. *IEEE Transactions on Learning Technologies*, 10(1), 17-29.

[13] Fritz, J. (2011). Classroom walls that talk: Using online course activity data of successful students to raise self-awareness of underperforming peers. *The Internet and Higher Education*, 14(2), 89-97.

[14] Tang, Z., von Seekamm, K., Colina, F. E., & Chen, L. (2025). Enhancing student retention with machine learning: A data-driven approach to predicting college student persistence. *Journal of College Student Retention*, 27(1), 145-168.

[15] Karalar, H., Kapucu, C., & Gürüler, H. (2021). Predicting students at risk of academic failure using ensemble model during pandemic in a distance learning system. *International Journal of Educational Technology in Higher Education*, 18, Article 63.

[16] Macfadyen, L. P., & Dawson, S. (2010). Mining LMS data to develop an "early warning system" for educators: A proof of concept. *Computers & Education*, 54(2), 588-599.

[17] Agudo-Peregrina, Á. F., Iglesias-Pradas, S., Conde-González, M. Á., & Hernández-García, Á. (2014). Can we predict success from log data in VLEs? Classification of interactions for learning analytics and their relation with performance in VLE-supported F2F and online learning. *Computers in Human Behavior*, 31, 542-550.

[18] Lakkaraju, H., Aguiar, E., Shan, C., Miller, D., Bhanpuri, N., Ghani, R., & Addison, K. L. (2015). A machine learning framework to identify students at risk of adverse academic outcomes. *Proceedings of the 21st ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 1909-1918.

[19] Sciencedirect. (2025). Predicting student retention: A comparative study of machine learning algorithms. *Decision Analytics Journal*, 14, Article 100456.

[20] Macfadyen, L. P., & Dawson, S. (2010). Mining LMS data to develop an "early warning system" for educators: A proof of concept. *Computers & Education*, 54(2), 588-599.

[21] Lakkaraju, H., Aguiar, E., Shan, C., Miller, D., Bhanpuri, N., Ghani, R., & Addison, K. L. (2015). A machine learning framework to identify students at risk of adverse academic outcomes. *Proceedings of the 21st ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 1909-1918.

[22] Mapademics. (2025). Predictive analytics for student retention: How AI is transforming higher education. White Paper Series on Educational Technology.

[23] Pardo, A., Jovanovic, J., Dawson, S., Gašević, D., & Mirriahi, N. (2019). Using learning analytics to scale the provision of personalised feedback. *British Journal of Educational Technology*, 50(1), 128-138.

[24] Yorke, M., & Longden, B. (2008). *The first-year experience of higher education in the UK*. Higher Education Academy.

[25] Strayhorn, T. L. (2018). *College students' sense of belonging: A key to educational success for all students* (2nd ed.). Routledge.

[26] Tempelaar, D., Rienties, B., & Nguyen, Q. (2020). Subjective data, objective data and the role of bias in predictive modelling: Lessons from a dispositional learning analytics application. *PLOS ONE*, 15(6), e0233977.

[27] Hosmer, D. W., Lemeshow, S., & Sturdivant, R. X. (2013). *Applied logistic regression* (3rd ed.). Wiley.

[28] Howard, E., Meehan, M., & Parnell, A. (2018). Contrasting prediction methods for early warning systems at undergraduate level. *The Internet and Higher Education*, 37, 66-75.

[29] Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5-32.

[30] Tang, Z., von Seekamm, K., Colina, F. E., & Chen, L. (2025). Enhancing student retention with machine learning: A data-driven approach to predicting college student persistence. *Journal of College Student Retention*, 27(1), 145-168.

[31] Araya Santiago, P. (2024). Student academic risk prediction [Software]. GitHub. https://github.com/pauloarayasantiago/student-academic-risk-prediction

[32] Hussain, M., Zhu, W., Zhang, W., Abidi, S. M. R., & Ali, S. (2019). Using machine learning to predict student difficulties from learning session data. *Artificial Intelligence Review*, 52(1), 381-407.

[33] Prenkaj, B., Velardi, P., Stilo, G., Distante, D., & Faralli, S. (2020). A survey of machine learning approaches for student dropout prediction in online courses. *ACM Computing Surveys*, 53(3), 1-34.

[34] National Student Clearinghouse Research Center. (2024). *Persistence and retention report 2024*. Herndon, VA: Author.

[35] Inside Higher Ed. (2025, June 26). First-year persistence continues slow climb from pandemic drop. Retrieved from https://www.insidehighered.com/

[36] CEUR Workshop Proceedings. (2019). Predictors and early warning systems in higher education: A literature review. *CEUR-WS.org*, Vol-2415.