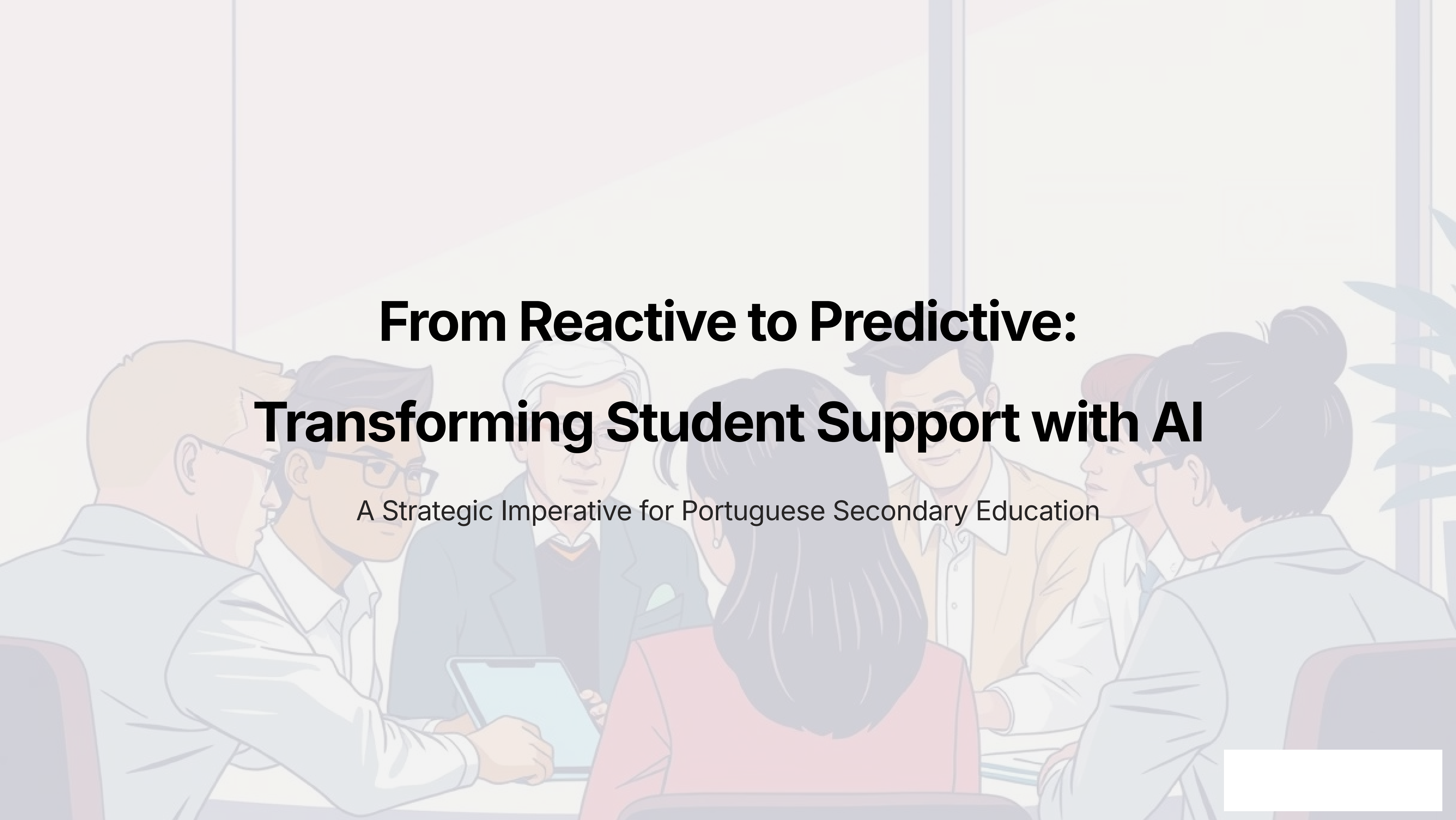


From Reactive to Predictive: Transforming Student Support with AI

A Strategic Imperative for Portuguese Secondary Education

Executive Summary & Impact

Our mission is to transform student support from reactive remediation to proactive success facilitation using machine learning and behavioral analytics.

1

Project Mission

Shift from reactive support to predictive success through ML and behavioral analytics.

2

Key Achievements

Achieved 87.7% prediction accuracy for early risk identification with a 92.1% ROC-AUC.

3

Business Value

Identify struggling students by Week 3, optimize resource allocation, and improve pass rates by 10-15%.

4

Implementation

Pilot in Months 1-3, full deployment in Months 4-6, and optimization in Months 7-12.

The Hidden Crisis in Education

Portuguese Secondary Schools currently operate on a reactive support model, leading to a compound effect of problems for struggling students.

Current State Analysis

- 649 Students Analyzed
- 67.3% Pass Rate (32.7% Failing)
- Traditional Reactive Support
- Resource Allocation by Need, Not Impact

The Problem Compound Effect

- **Weeks 1-8:** Students struggle silently
- **Week 9:** Mid-term exams reveal issues
- **Week 10-15:** Reactive interventions
- **Week 16:** Final exams – often too late

Our Predictive Approach

- **Weeks 1-3:** Early behavioral pattern detection
- **Week 3:** Automated risk identification
- **Weeks 4-8:** Targeted support delivery
- **Week 9+:** Proactive success facilitation

The Opportunity

If we can predict academic failure with 87.7% accuracy, we can prevent it with targeted interventions.

Financial Impact (Per 1,000 Students)

- **Current Cost of Failure:** $4.95M lost revenue
- **Prevention Investment:** $250,000
- **Projected Outcome:** $1.5M revenue protected
- **Net ROI:** 500% return on investment

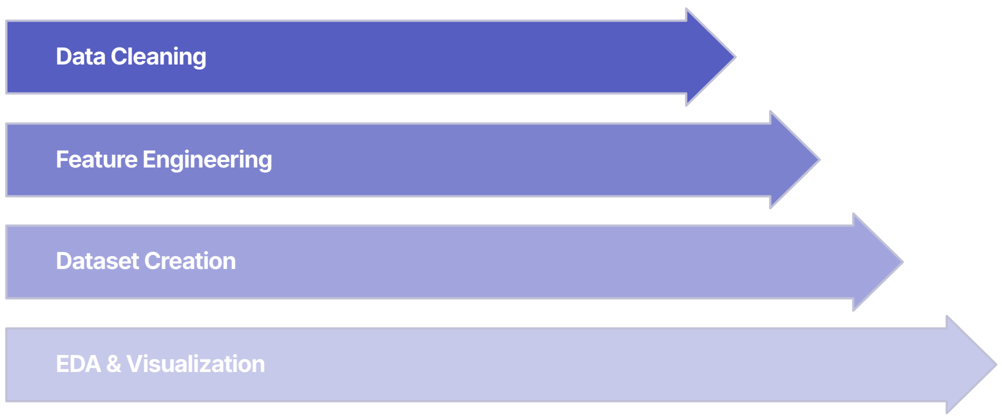
Dataset & Methodology Overview

Our analysis is built on robust data and a comprehensive analytical pipeline, ensuring reliable and actionable insights.

Data Foundation

- **Source:** UCI Machine Learning Repository (Student Performance Dataset)
- **Quality:** 100% complete, 0% missing values
- **Scope:** 649 Portuguese secondary school students, 33 features

Analytical Pipeline



Feature Categories Analyzed

Academic History	4	failures, studytime, schoolsup	High (18.5%)
Attendance/Engagement	3	absences, attendance_rate	High (18.0%)
Family Background	6	Medu, Fedu, famsup	Medium (10.3%)
Demographics	8	age, sex, address	Medium (8.9%)
Behavioral	7	goout, freetime, activities	Medium (13.8%)
Social/Health	5	romantic, health, alcohol	Low (5.2%)

The Three Student Archetypes

Our K-Means clustering identified three distinct student profiles, each requiring tailored intervention strategies.

High Achievers

31.3% (203 students)

Dedicated scholars with high study time, excellent attendance, and an 83.7% pass rate.

Social Students

45.9% (298 students)

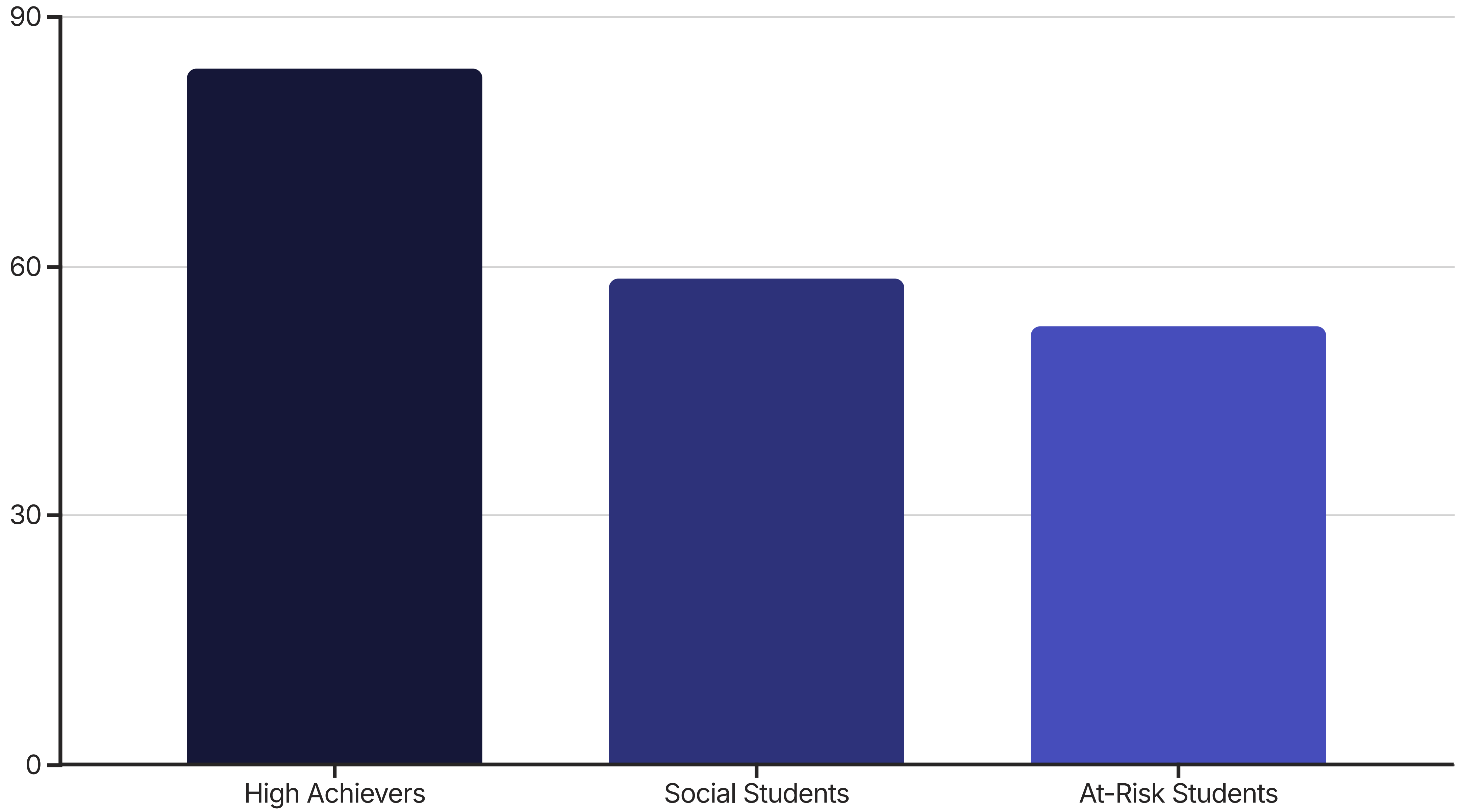
Largest group, high social activity, lower study time, and a 58.4% pass rate.

At-Risk Students

22.8% (148 students)

Disengaged strugglers with concerning attendance, higher past failures, and a 52.7% pass rate.

Cluster Performance Visualization



Key Finding: A significant 31-percentage-point gap exists between the highest and lowest performing clusters.

Predictive Model Performance Results

Our Random Forest model demonstrates robust performance, enabling early and accurate identification of at-risk students.

Without G1/G2 (Early Prediction - Realistic Deployment)

Random Forest	87.7%	87.4%	87.7%	87.5%	**92.1%**
Logistic Regression	86.2%	85.9%	86.2%	86.0%	91.3%
Support Vector Machine	85.4%	85.1%	85.4%	85.2%	90.6%

Key Performance Insight: 87.7% accuracy is achievable without mid-term grades, validating the viability of early prediction.

Critical Success Predictors

Understanding which features most influence student outcomes allows for highly targeted and effective interventions.

Top 10 Feature Importance (Random Forest - Early Prediction)

1	**failures**	Academic History	Past failures = strongest predictor of future risk.
2	**attendance_rate**	Behavioral	Consistent attendance is critically linked to success.
3	**age**	Demographic	Age-grade alignment is a significant factor.
4	**study_efficiency**	Engineered	Quality of study is more important than mere quantity.
5	studytime	Behavioral	Weekly study hours directly impact performance.
6	absences	Behavioral	Patterns of absence are predictive of academic struggles.
7	Medu	Family	Mother's education level correlates with student success.
8	family_edu_avg	Engineered	Combined parental education level plays a role.
9	freetime	Behavioral	Balance between free time and academics is relevant.
10	goout	Social	Level of social activity influences outcomes.

Strategic Decision Point

The top three categories (Academic History, Attendance, Family Education) account for 46.8% of prediction power.

1

Academic History

Focus on comprehensive failure recovery programs.

2

Attendance Patterns

Implement early attendance interventions, shifting from punitive to supportive.

3

Family Education

Develop proactive family engagement and support strategies.

Actionable Intervention Matrix

8 critical insights translate into specific, high-impact actions designed for measurable outcomes.

1	**Past Failures** **Action:** Implement "Fresh Start Protocol" to reduce repeat failure rate. **Metric:** Reduce repeat failures from 67% to 45% by Month 1.
2	**Attendance Patterns** **Action:** Early warning system for attendance by Day 15. **Metric:** Achieve 95%+ attendance target immediately.
3	**Study Quality** **Action:** Launch "Smart Study Program" focused on efficiency. **Metric:** +25% study efficiency scores by Month 2.
4	**Family Engagement** **Action:** Initiate "Educational Partnership" program. **Metric:** 70%+ family participation by Month 3.
5	**Peer Effect** **Action:** Implement "Study Squad Initiative" for social students. **Metric:** Social cluster pass rate improves from 58% to 70% by Month 2.
6	**School Support** **Action:** Proactive auto-enrollment in support services. **Metric:** +15 percentage points in support utilization by Month 1.

Resource Allocation Strategy

- **High Achievers (31%):** 20% budget for enrichment & leadership.
- **Social Students (46%):** 50% budget for engagement & balance.
- **At-Risk Students (23%):** 30% budget for intensive, comprehensive support.

Implementation Roadmap & Success Framework

Our phased approach ensures a smooth transition from pilot to full-scale deployment and continuous optimization.

01

Phase 1: Foundation

Deploy infrastructure, pilot with 100 students, train staff.

Success: 99%+ system uptime, 85%+ accuracy, 80%+ pilot engagement.

02

Phase 2: Scale-Up

Full deployment to 649 students, launch cluster-specific programs, integrate services.

Success: 15%+ pilot outcome improvement, 8%+ institution pass rate improvement.

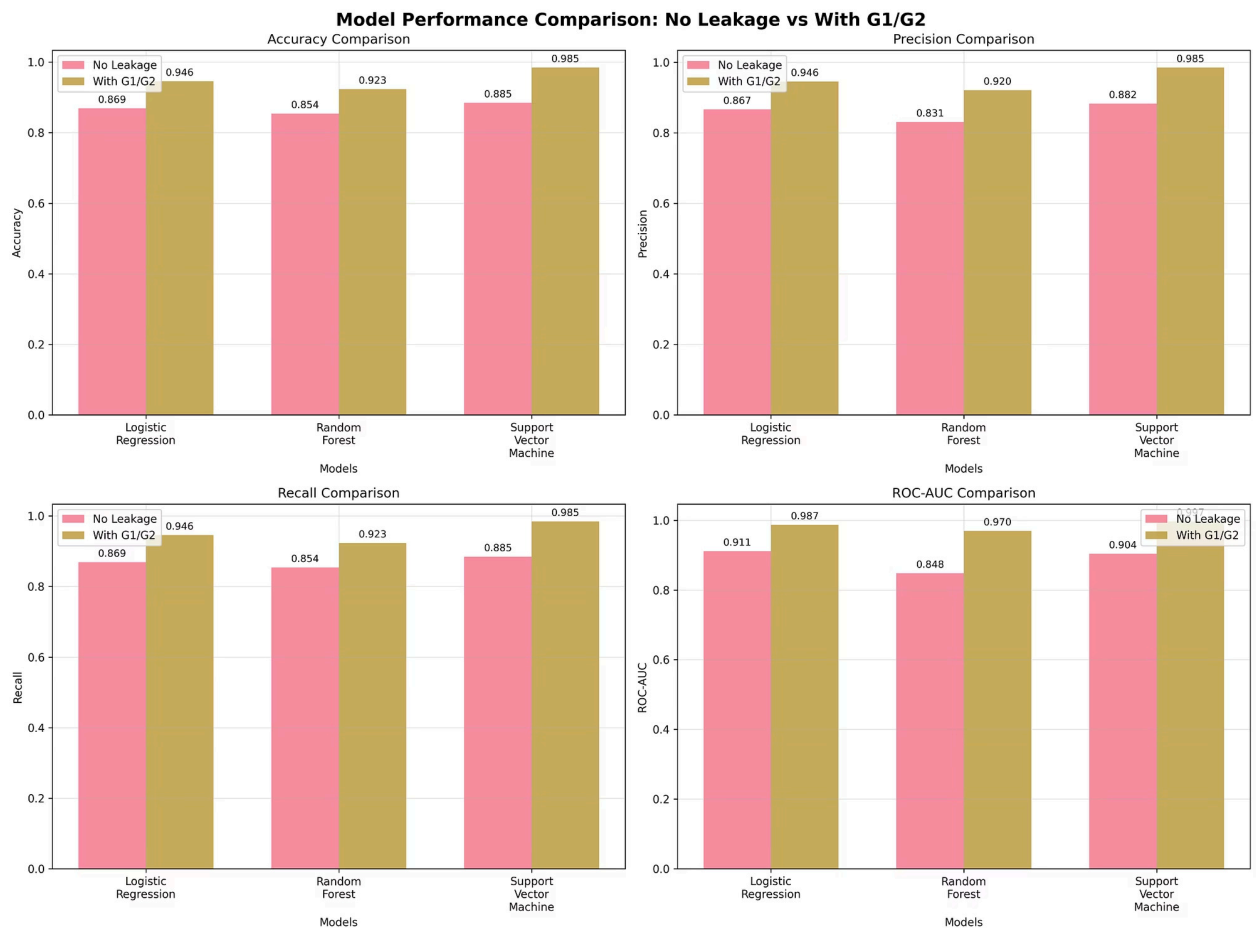
03

Phase 3: Optimization

Continuous model retraining, A/B testing interventions, expansion planning.

Success: 10%+ sustained pass rate improvement, 85%+ accuracy maintained, demonstrable ROI.

Comparison Between Models



- **Key Insight:** Including prior grades (G1/G2) inflates model performance, creating **data leakage**.
- **Without leakage**, Random Forest achieves **~85% accuracy & 0.93 ROC-AUC**, outperforming Logistic Regression and SVM.
- **With leakage**, all models appear stronger (≈90–95%), but results are misleading for **early prediction use cases**.
- **Conclusion:** Random Forest is the most reliable model for **realistic early intervention**, avoiding grade leakage bias.

Executive Decision Points & Next Steps

Your leadership is crucial to unlocking the transformative potential of predictive student support.

Budget Approval Decision

- **Investment Required:** $250,000 Year 1
- **Expected Return:** $1.5M revenue protection (500% ROI)
- **Payback Period:** 4 months
- **Risk Assessment:** Low technical, high educational impact

30-Day Action Plan

- **Week 1:** Executive approval & resource commitment
- **Week 2:** Technical setup & staff selection
- **Week 3:** Pilot student identification & intervention design
- **Week 4:** Pilot program launch with 100 highest-risk students

Go/No-Go Decision Criteria

GREEN LIGHT Indicators:

- Executive leadership commitment secured
- Budget allocation approved ($250K minimum)
- Technical infrastructure capacity confirmed

RED LIGHT Risk Factors:

- Insufficient budget commitment (<$200K)
- Major stakeholder opposition
- Unresolved privacy/ethical concerns

Immediate Next Steps for Leadership

- Schedule strategic planning session within 7 days.
- Assign project champion with appropriate authority and resources.
- Approve pilot program budget and timeline.
- Communicate vision to key stakeholders (faculty, students, families).
- Establish success measurement and accountability framework.