

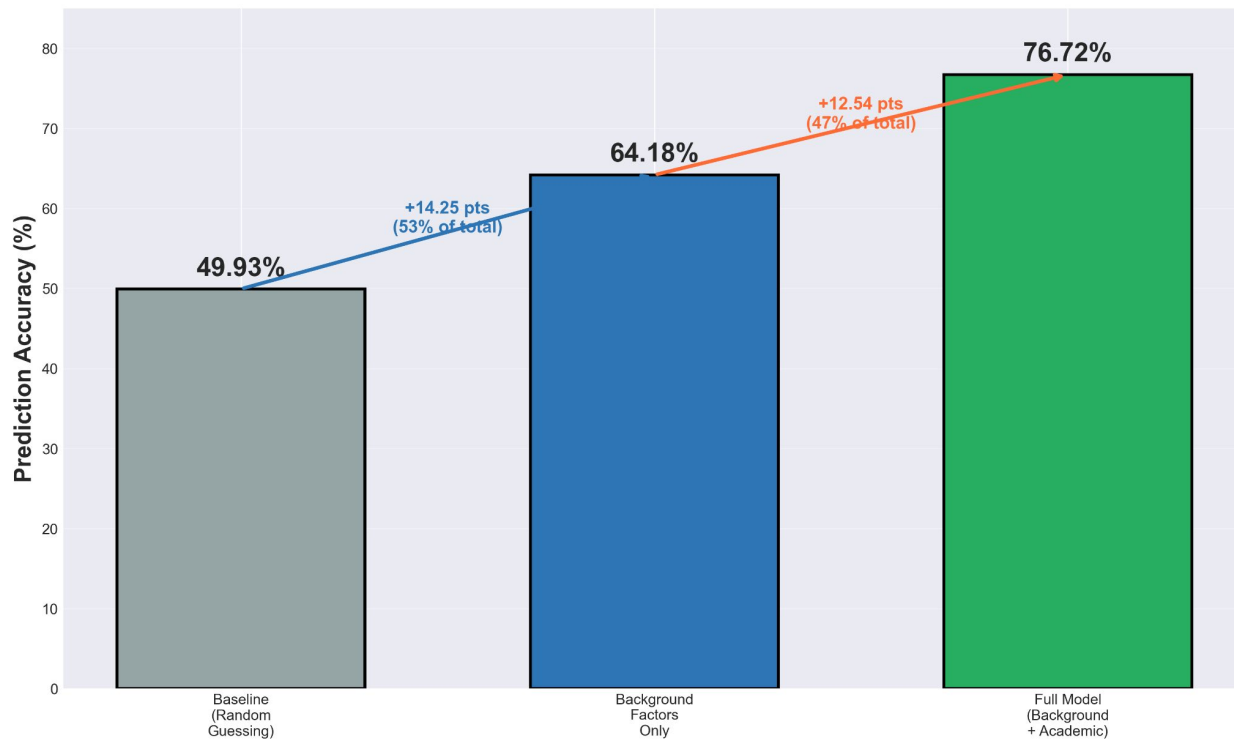
What Really Predicts Student Dropout?

(53% ≠ Grades)

Sue Sue | CSC-466

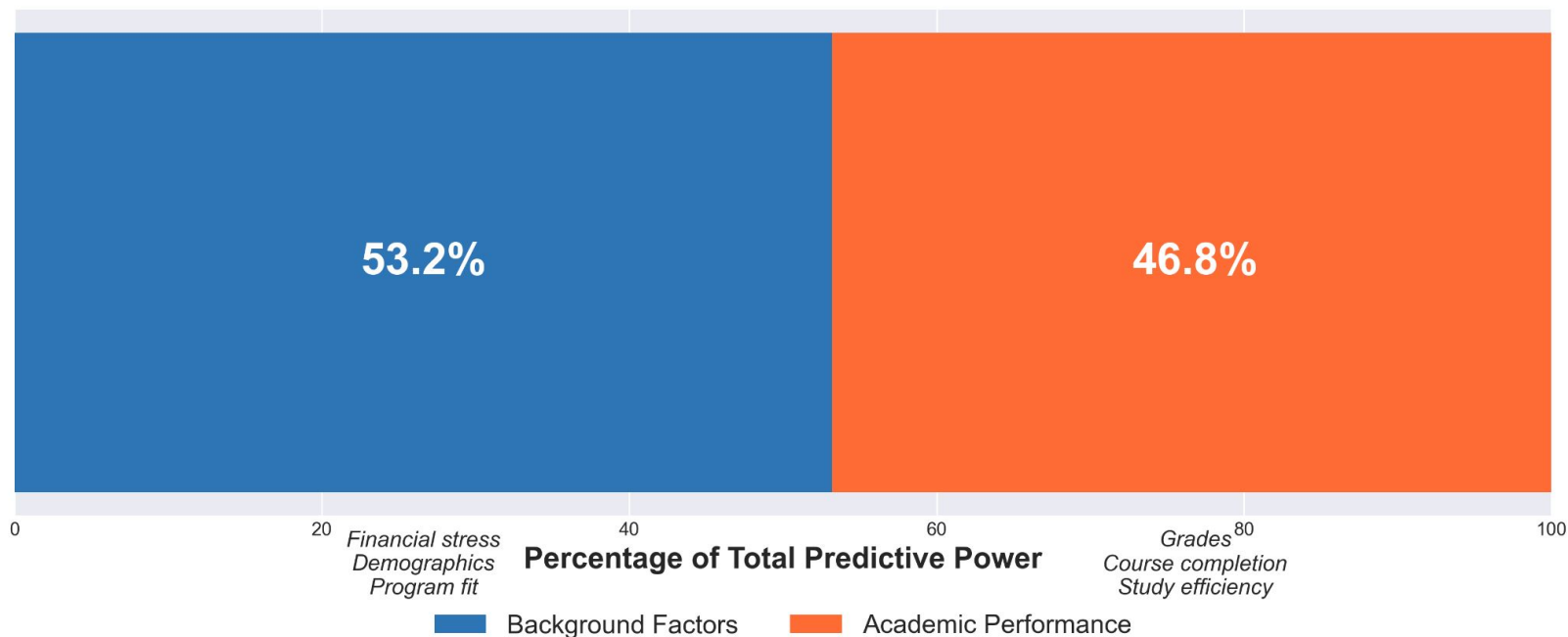
The 53/47 Split Finding

Building Predictive Power: From Baseline to Full Model
Background Factors Contribute 53% of Total Improvement



The 53/47 Split Finding

Where Does Predictive Power Come From?
53% Background vs. 47% Academic



The 53/47 Split Finding

Background Factors: 53.2%

- Financial status (tuition, scholarships, debt)
- Demographics (age, gender, family background)
- Program selection and prior qualifications
- Available at admission - DAY ONE

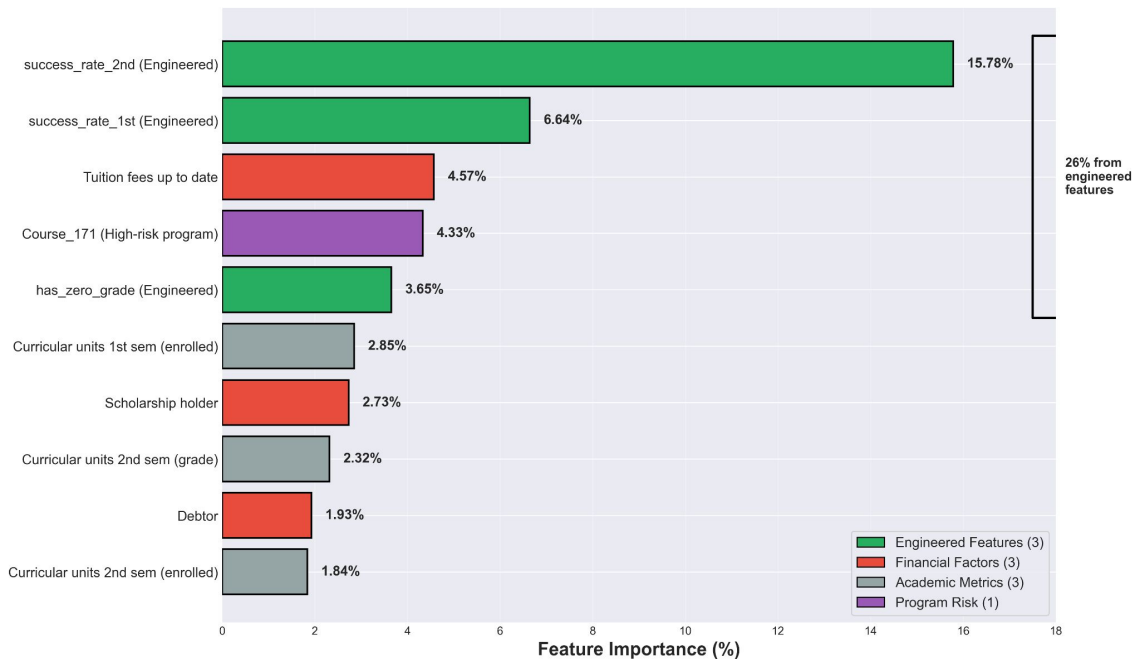
Academic Performance: 46.8%

- Grades, course completion, study efficiency
- Available after Semester 2

Implication: Most risk exists BEFORE academic struggles

Evidence - What Actually Predicts Dropout

What Actually Predicts Student Dropout?
Top 10 Features Ranked by Importance



Engineered Features (26% total):

- Success_rate_2nd - 15.78%
- Success_rate_1st - 6.64%
- Has_zero_grade - 3.65%

Financial Factors (9.2% total):

- Tuition fees up to date - 4.57%
- Scholarship holder - 2.73%
- Debtor status - 1.93%

Key Finding: Raw grades rank only #8
- Context matters more than numbers alone

The Solution - Two-Stage Intervention System

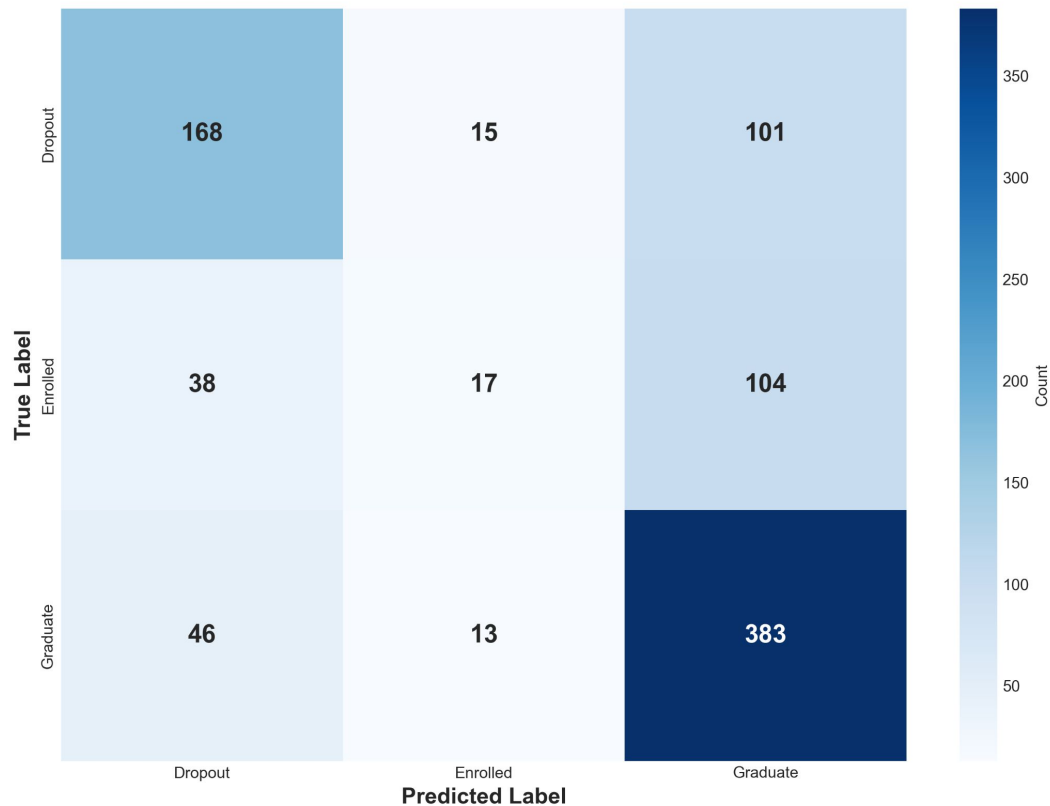
STAGE 1 - Admission Screening

- Background model: 64% accuracy
- Detects 59% of future dropouts (168/284)
- Timeline: Before classes start
- Interventions
 - Financial aid counseling
 - Emergency grants for at-risk students
 - Enhanced orientation programs
 - Program-specific mentoring

STAGE 2 - Academic Monitoring

- Full model: 77% accuracy
- Detects 76% of dropouts (217/284)
- Timeline: After semester 2 grades
- Interventions:
 - Mandatory tutoring
 - Academic probation protocols
 - Course load adjustments
 - Intensive advising

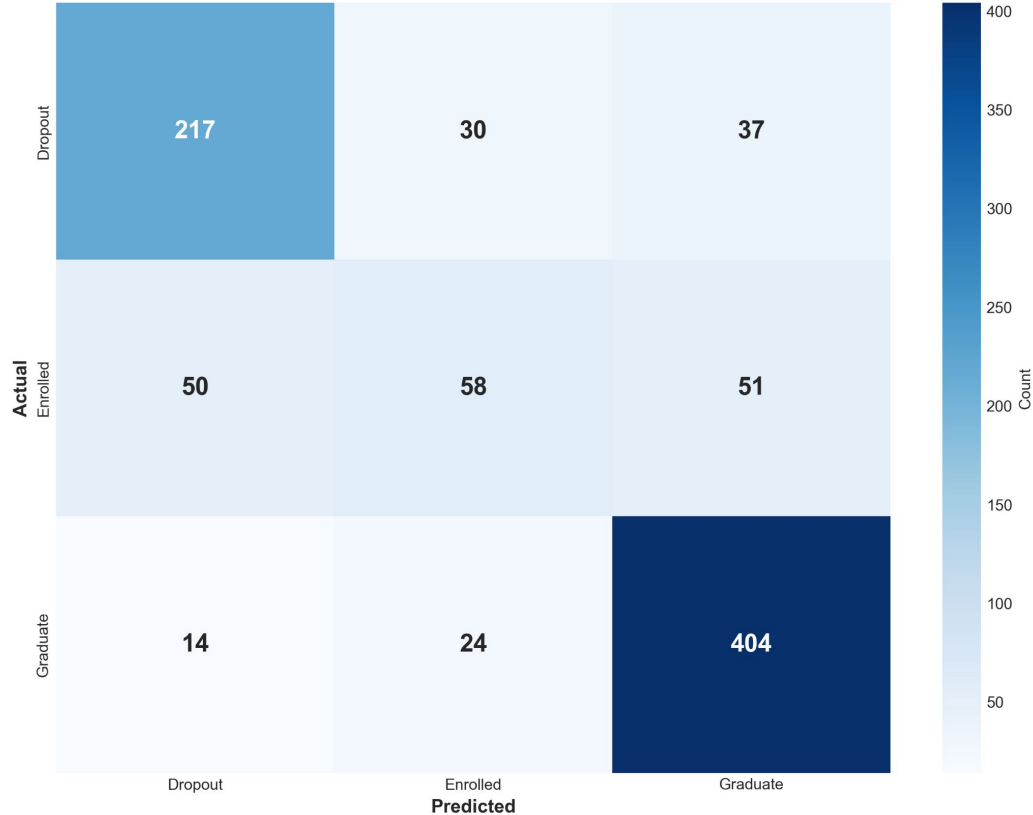
Confusion Matrix - Background-Only Model (XGBoost)
Test Accuracy: 64.18%



Background-Only (At Admission):

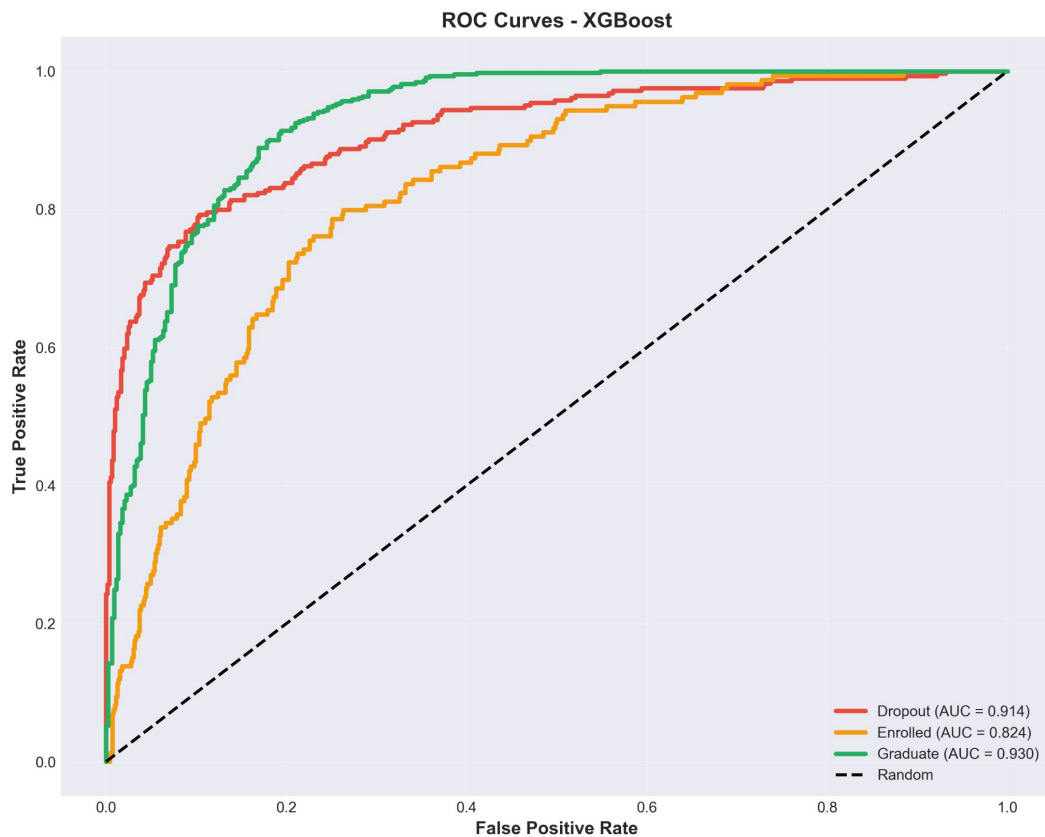
- 59% Dropout recall (168/284)
- 87% Graduate recall (383/442)
- 11%% Enrolled recall (17/159)

Confusion Matrix - Full Model (XGBoost)
Test Accuracy: 76.72%



Full Model (After Semester 2):

- 76% Dropout recall (217/284)
- 91% Graduate recall (404/442)
- 36% Enrolled recall (58/159)



Discrimination Ability (ROC-AUC Scores)

- Graduate: 0.93 AUC
- Dropout: 0.91 AUC
- Enrolled: 0.82 AUC
- Macro Average: 0.87

Impact and Final Takeaway

Key Results

- 76.72% prediction accuracy (XGBoost)
- 217/284 dropout cases identified (76% recall)
- 53/47 split challenges conventional knowledge

Implications

- Don't wait for academic failure to intervene
- Address structural barriers at enrollment
- Financial stress + program fit = actionable targets