Predicting School Closure

Can we Identify Schools at Risk for Closure by Performance and Other Characteristics?

Label: School Will Close within 5 Years

Features:

School Characteristics

- Geographic Region
- Elementary, Middle, or High
- Urban, Suburban, or Rural
- Magnet School (yes/ no)
- Charter School (yes/ no)
- Shared Time School (yes/ no)

Student Demographics

- Total Enrollment
- Percent Minority
- Percent Free/ Reduced Price Lunch
- Students Eligible for Title I Funding (all, some, or none)

School Performance

- Student-Teacher Ratio
- Math Percentile Rank
- ELA Percentile Rank

Classifier Evaluation

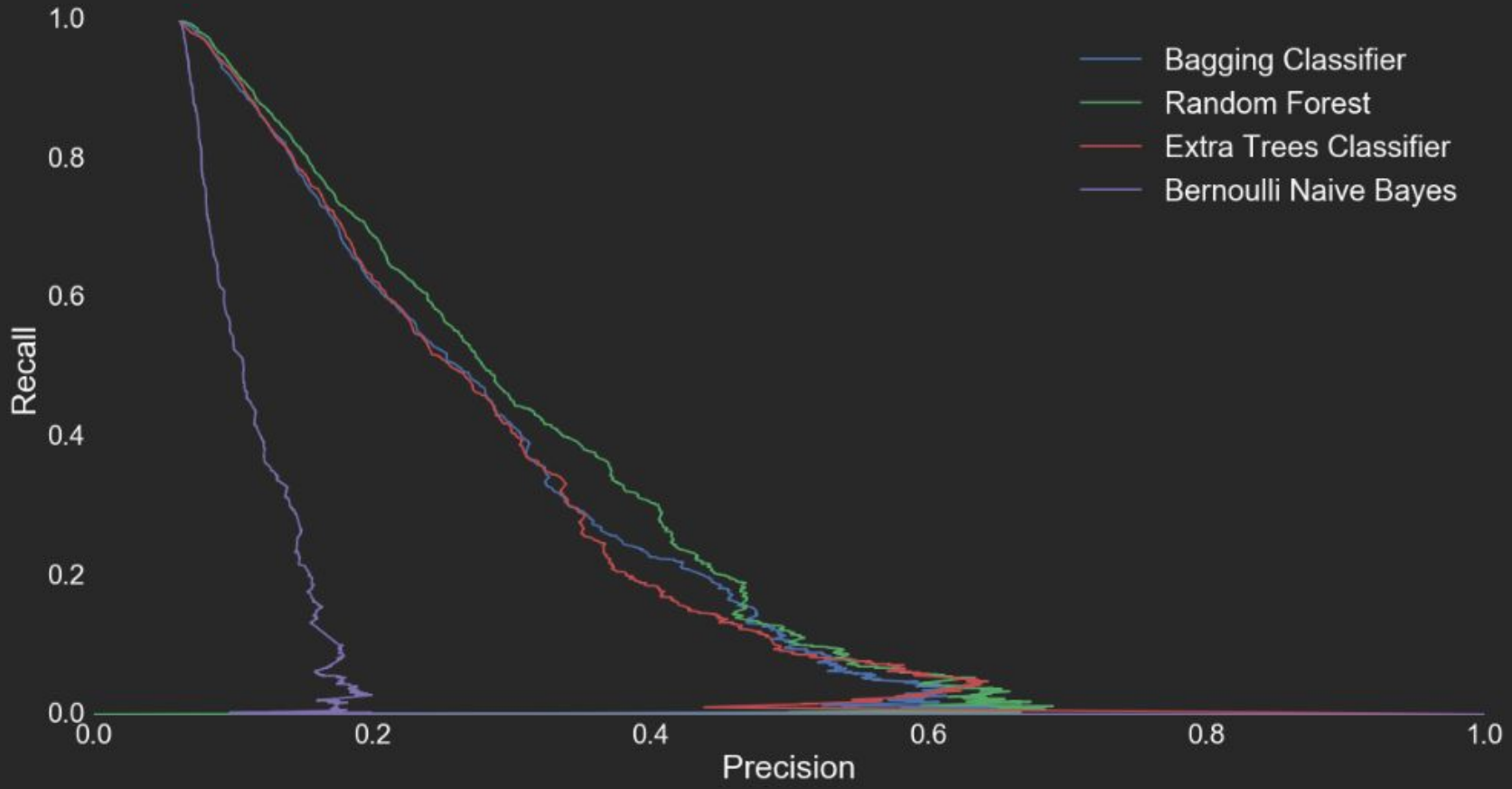
Recall

- Proportion of "At Risk" Schools Identified.
- Higher Recall = More *Effective* Policy

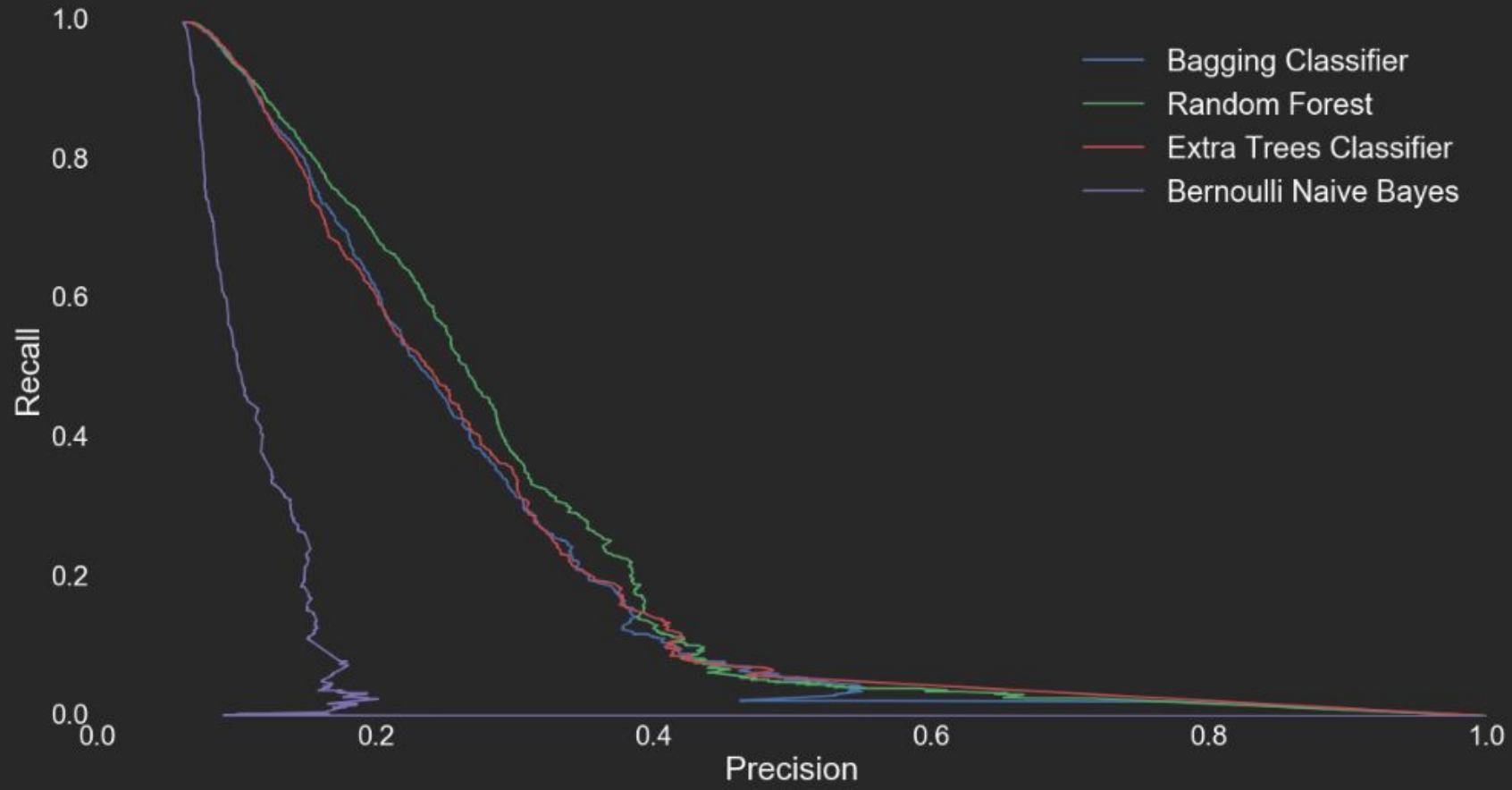
Precision

- Proportion of Identified Schools Actually at High Risk for Closure
- Higher Precision = More *Cost-Efficient* Policy

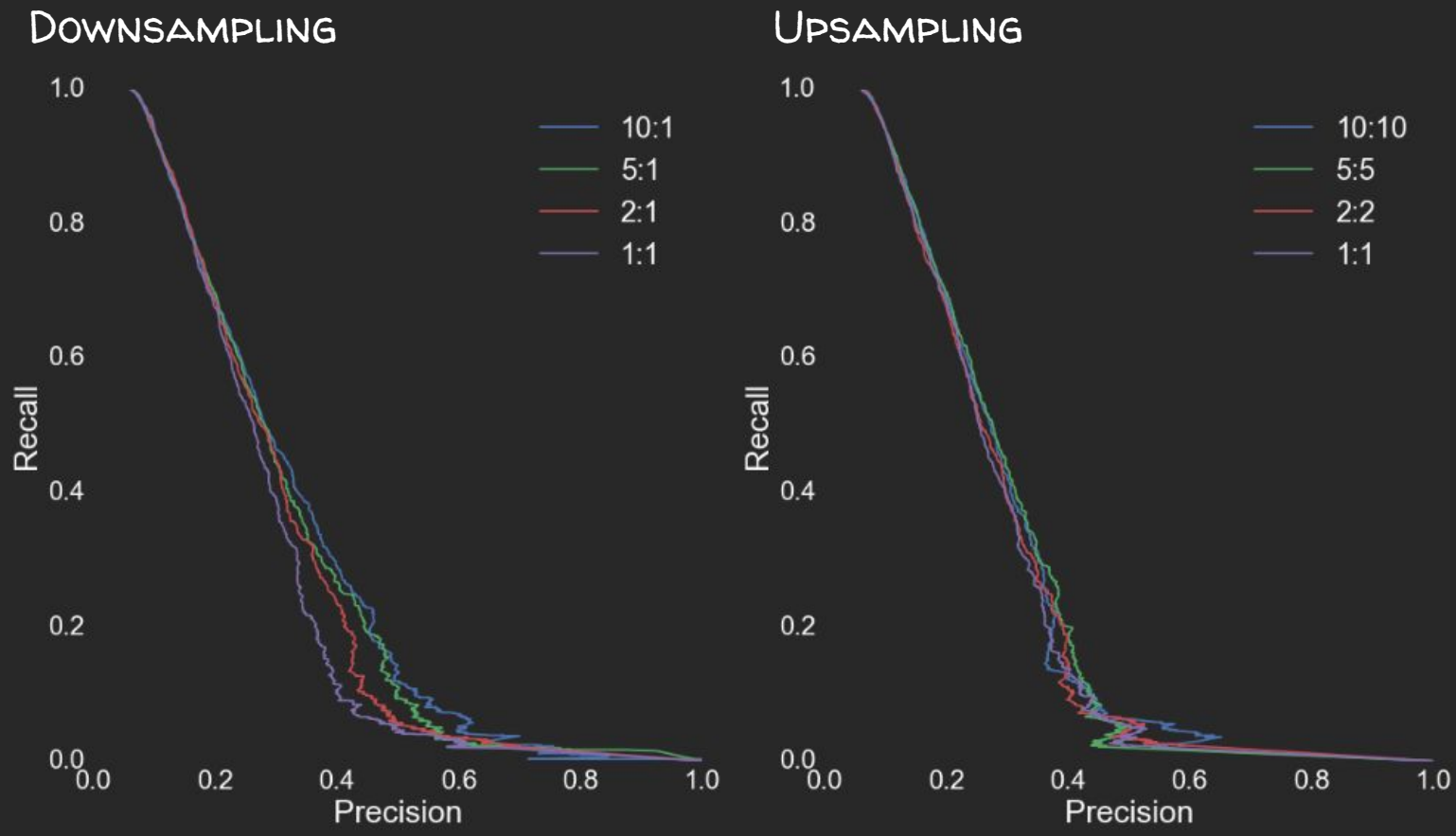
Classifier Precision-Recall Curves (on test set)



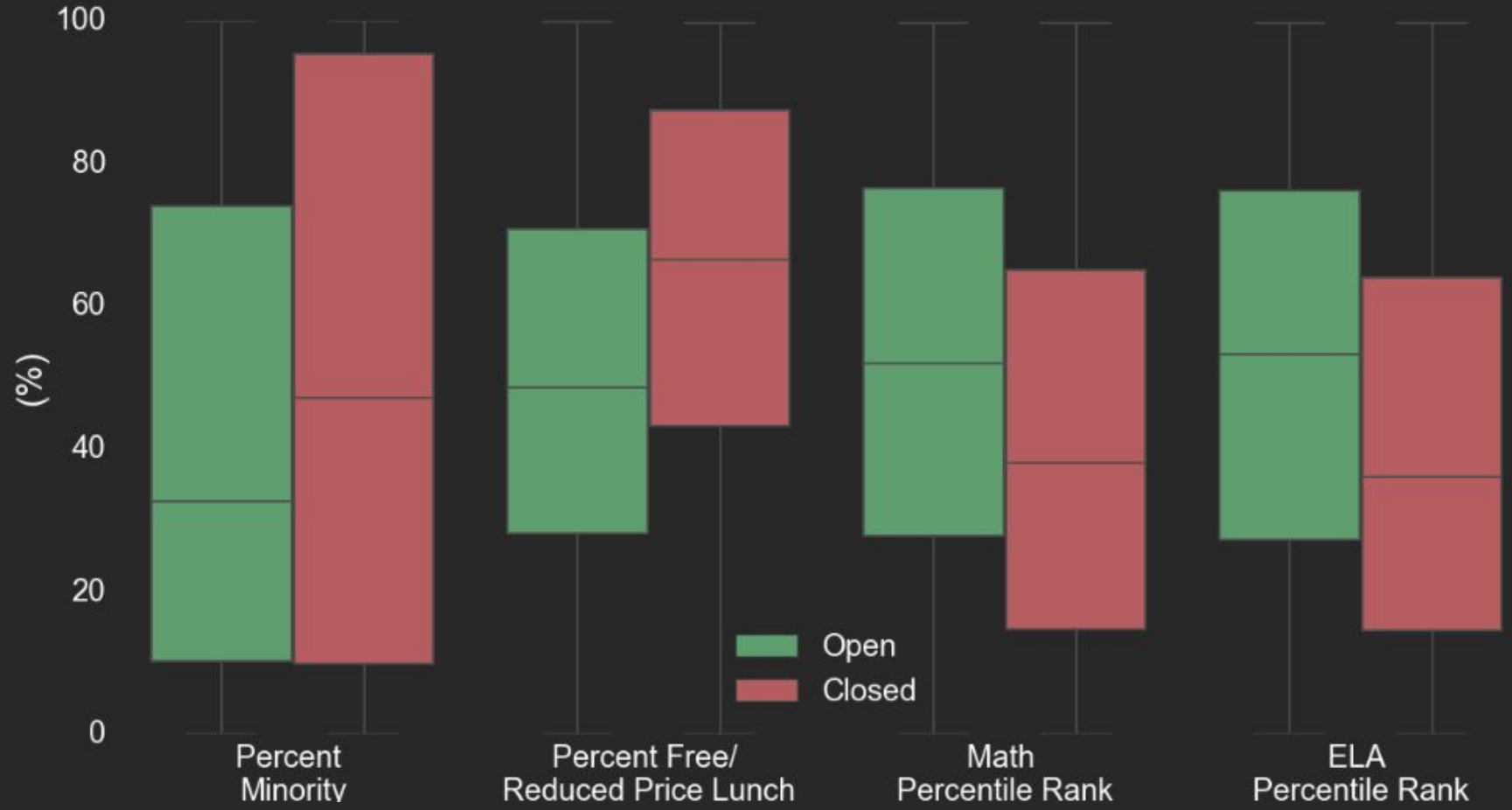
(2): 1:1 Downsampling



(3): Random Forest with Variable Downsampling & Upsampling

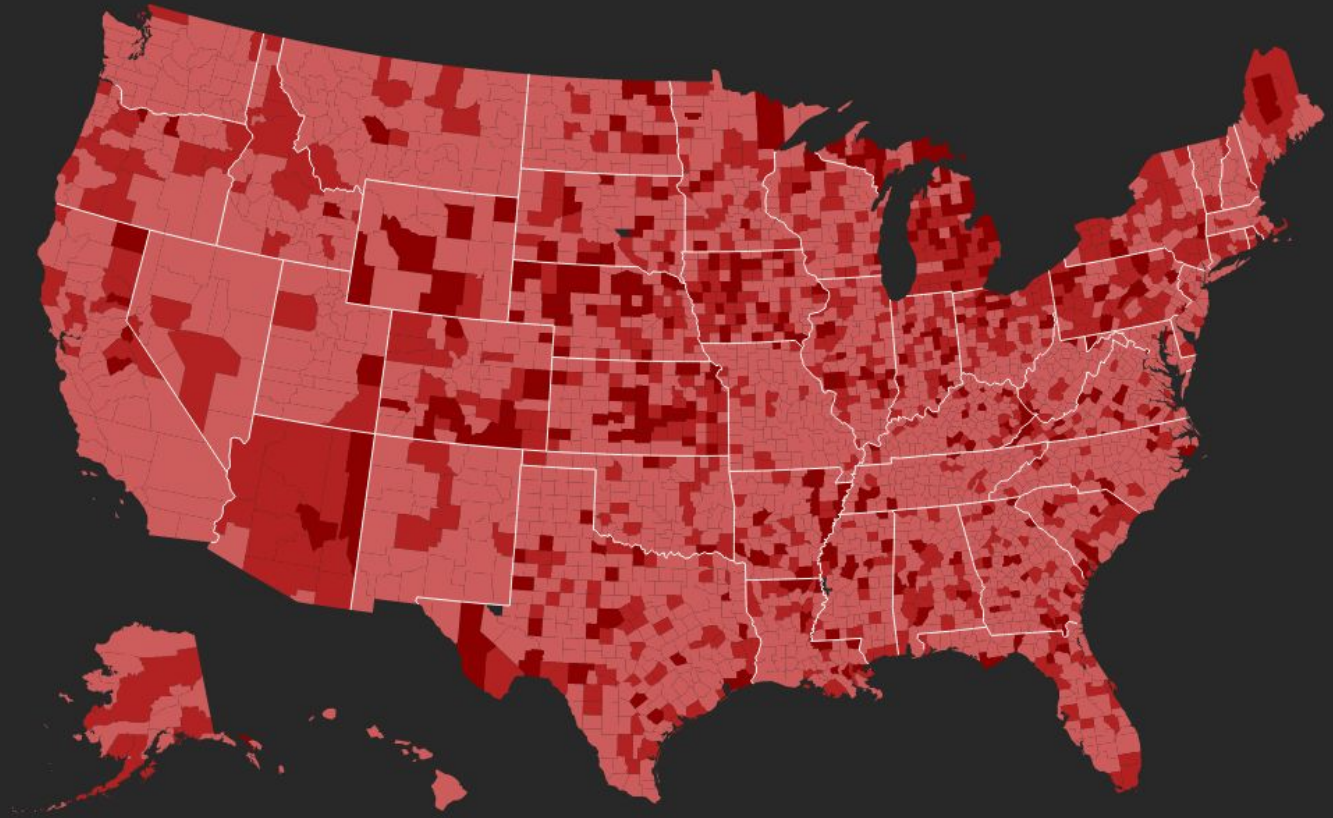


Characteristics of Open vs. Closed Schools



School Closures 2010-15

(Proportion of Closed Schools by County)



Insights

1. School characteristics *alone* insufficient to predict risk of closure.

2. Poor model performance cannot be explained purely by class imbalance.