
High School Graduation Rates

Background

- We will build a classification model which classifies high schools as either having a low or normal graduation rate for the United States in the year 2015.
 - The distinction between low and normal graduation rates is based on the federal government's standard that those schools with less than two thirds of class graduating are low graduation rate schools.
- Limitations:
 - Our data is limited to only those schools which reported their graduation rates
 - We looked at one year of data choosing 2015 as this was the year with the most complete reporting at the school and district level

Business Problem

- Predict which schools have low (66% and below) and which schools have normal (67% and above) high school graduation rates.
- Identify which characteristics are the best indicator of low and normal graduation rates so that school districts know where to focus resources when attempting to increase high school graduation rates.

What data did we use?

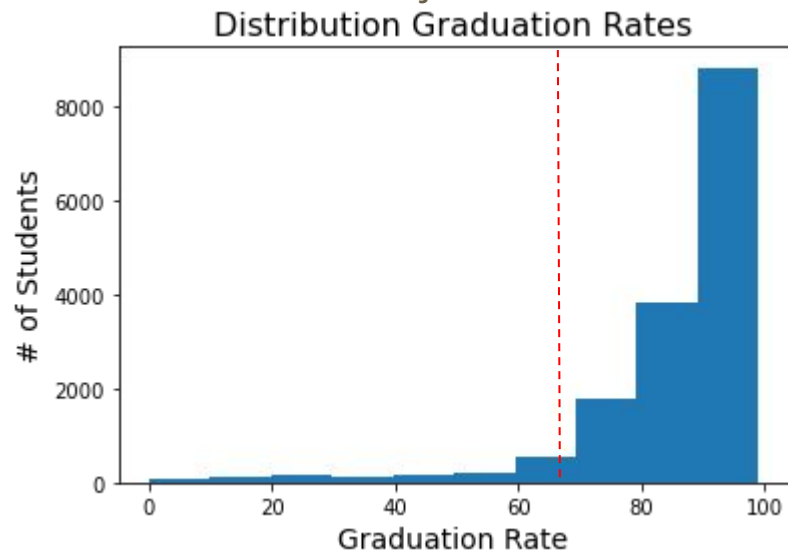
- Federal data on U.S. schools:
 - Every year, the federal government releases large amounts of data on U.S. schools, school districts, and colleges.
 - This information is scattered across multiple datasets that are often difficult to access and use.
- Urban Institute (<https://educationdata.urban.org/documentation/index.html>):
 - The Urban Institute Organizes and consolidates this data to make it easier to combine data from different reporting sources

Data Fields

- **School level data includes:**
 - school location
 - degree of urbanization
 - school size
 - # of students proficient in math and reading assessments
 - # of allegations for harassment/bullying
 - # of students enrolled in specific subjects
 - # of disciplinary actions taken
 - # of students participation in ACT/SAT tests
- **District level data includes:**
 - type and amount of capital outlays
 - debt outstanding at the end of the fiscal year
 - district expenditures and revenue
 - employee benefits
 - local state and federal revenue
 - salary for instruction

Distribution of National Graduation Rates

- Graduation rates are skewed towards low graduation rates.
- Low graduation rate high schools are the minority class making up less than 10% of all high schools.
- We will need to adjust for this imbalance in the modeling stage

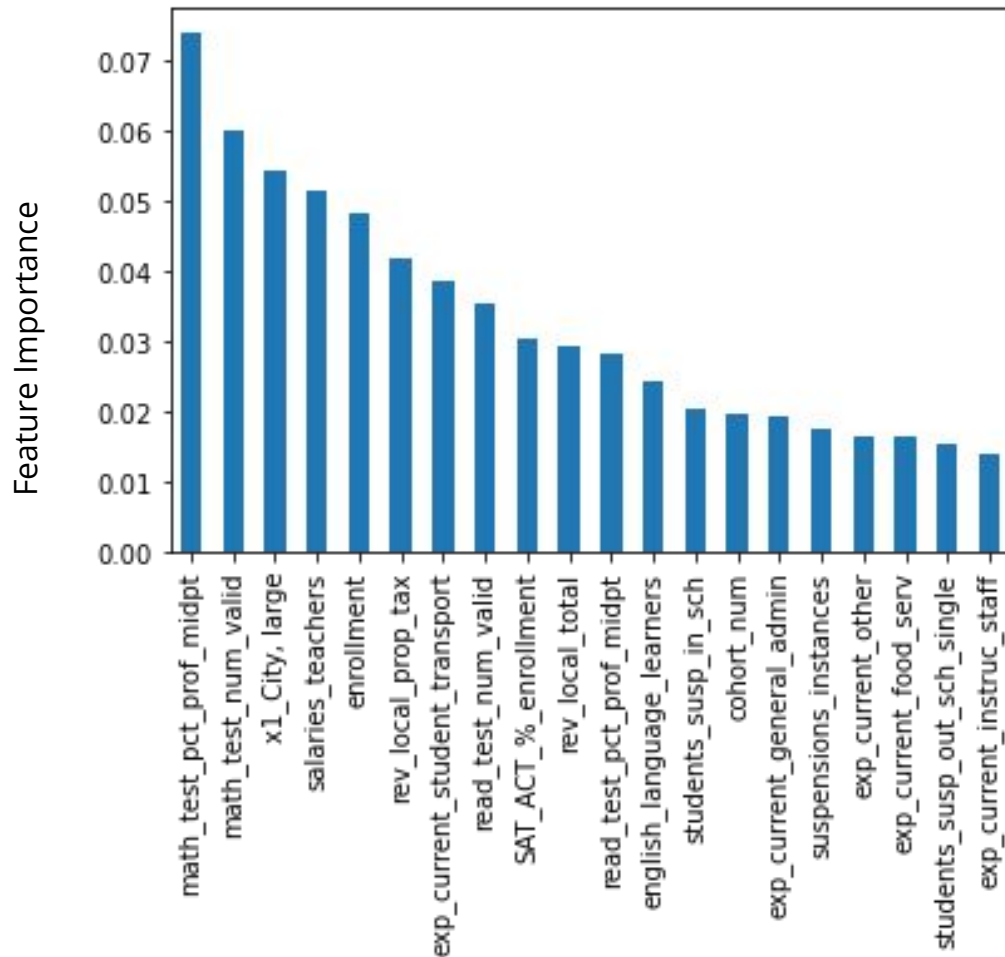


	Graduation Rate	% of Schools
Low Graduation Rate	0-66	7%
Normal Graduation Rate	67-100	93%

Feature Importance

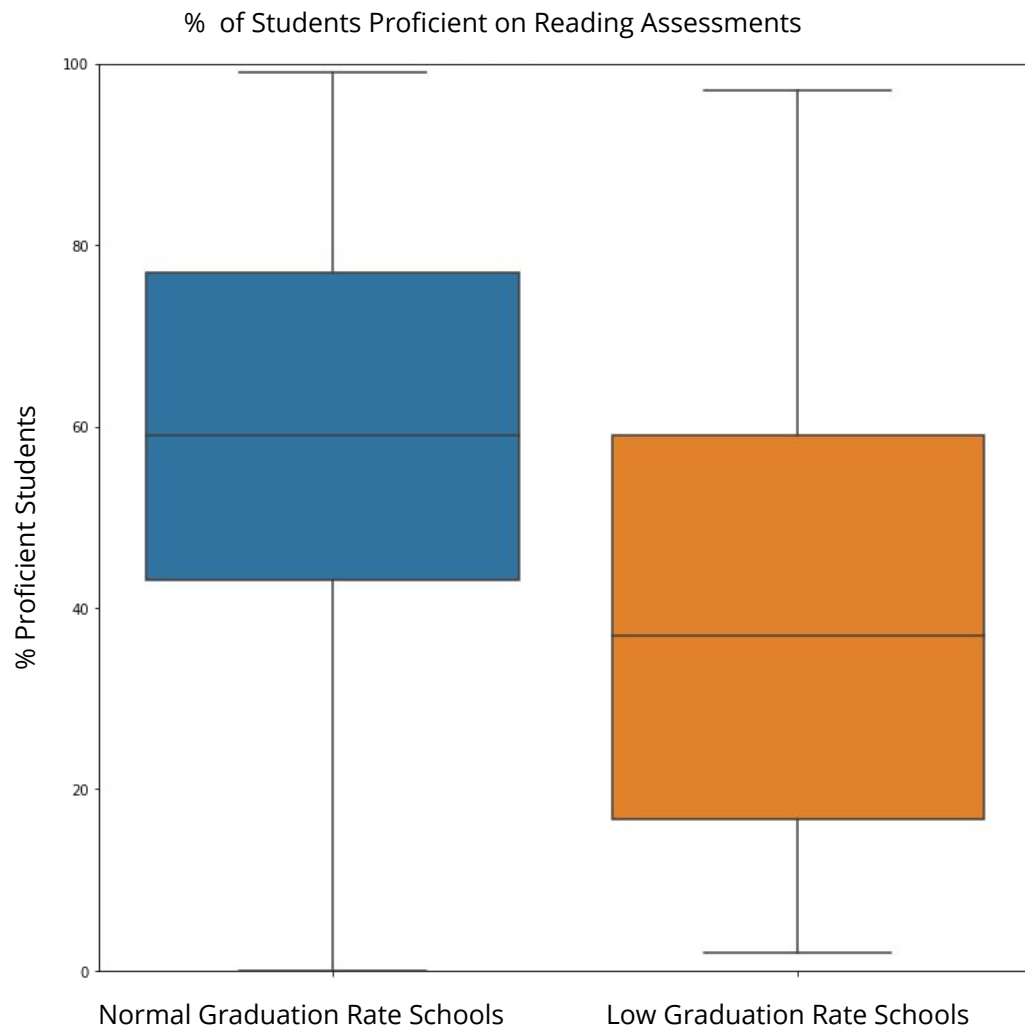
We will now take a closer look at how some of these top features are indicating for our model whether a high school should be predicted as a low or normal graduation rate.

Top 20 Features for Modeling



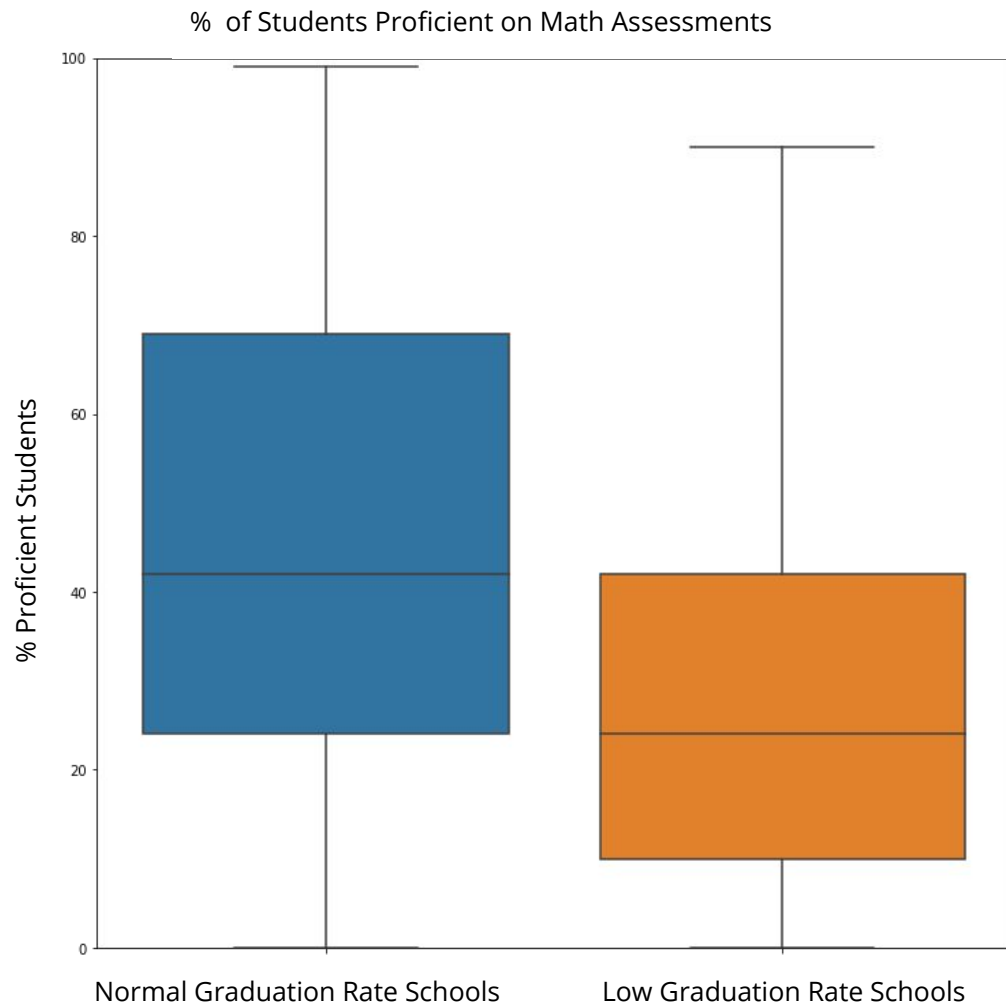
Reading

Higher reading test proficiency was associated with normal high school graduation rates



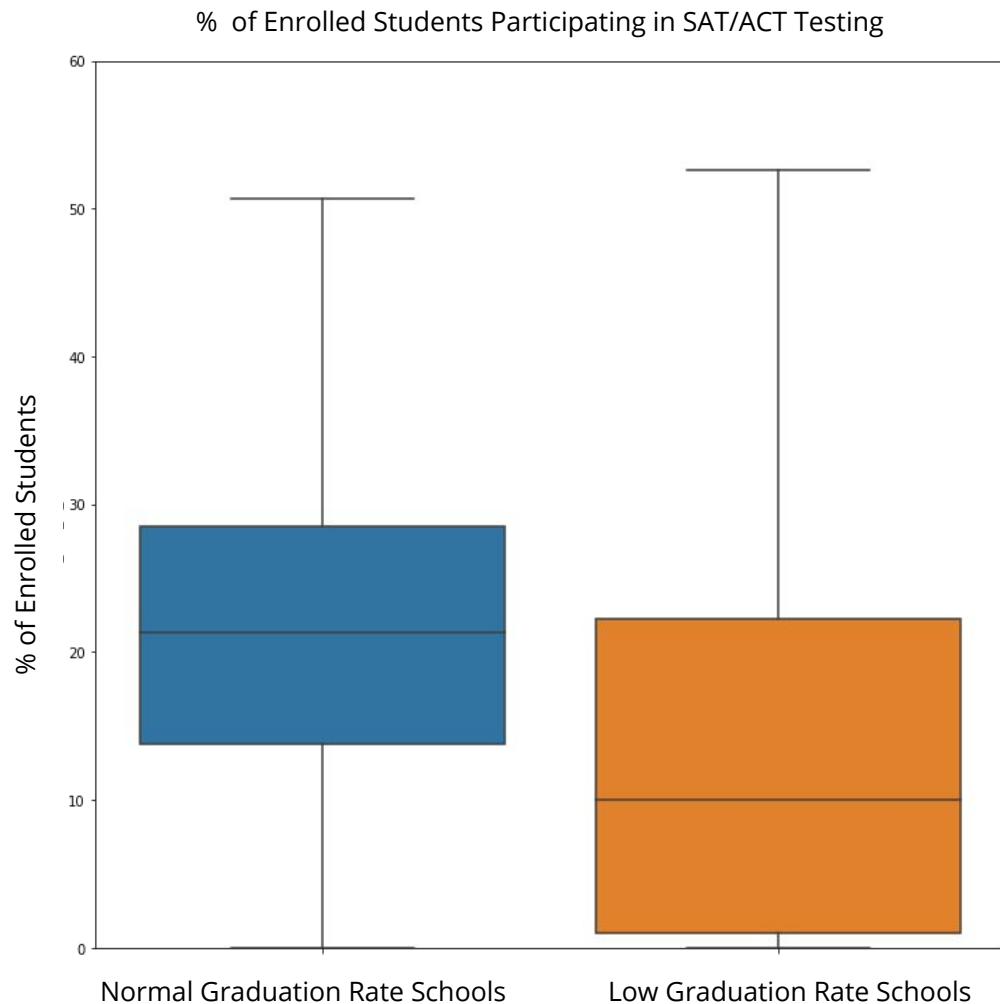
Mathematics

Higher math test proficiency was associated with normal high school graduation rates



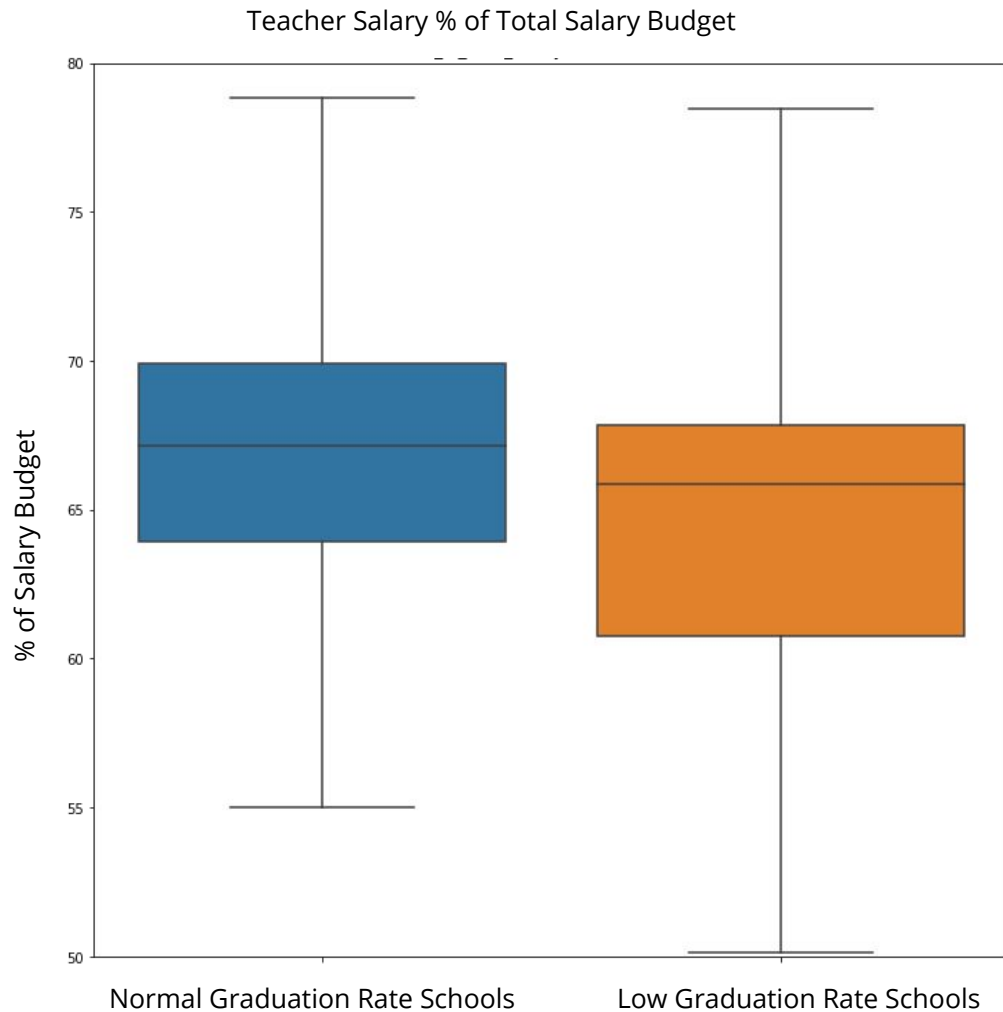
ACT/SAT Participation

Higher percent of ACT/SAT participation was associated with normal high school graduation rates



Teacher Salaries

Higher proportion of total salary budget devoted to teacher salary was associated with normal high school graduation rates



Conclusions

- Modeling:
 - struggles with overfitting
 - better at classifying the majority class (normal graduation rate)
 - Recall for normal grad rate was .99
 - Recall for low grad rate was .51
 - high enough performance that we can make recommendations based on feature performance
- Some important features for our model's classification were:
 - % of students who completed math and reading tests and received a proficient grade
 - % of enrolled students participating in ACT/SAT tests
 - % of Salary budget devoted to Teacher Salaries

Recommendations

- Mathematics and Reading:
 - Focus resources on providing extra support for mathematics and reading.
 - School districts should offer remedial opportunities, additional tutoring, and testing practice in order to improve math and reading scores.

Recommendations

- ACT/SAT:
 - As with many education metrics ACT/SAT participation likely also has a relationship with school funding and other types of factors
 - ACT/SAT prep can be a motivating factor for students to complete high school
- Teacher Salaries % of Salary Budget:
 - This feature may be confounded by other factors. For example a school with less overall funding likely has to devote more salary to maintenance workers if they are unable to receive funding for newer facilities.
 - We do not recommend making teacher salaries a higher proportion in budgeting plans as this indicator is likely confounded by other financial factors.

Future Work

- Perform feature selection steps to improve overfit of the modeling.
- Investigate adding new features from the Urban Institute reporting platform.
- Engineering new features with the data we already have on hand.
 - Engineering more features to make them proportional to the overall population of the graduating class
- Investigate which types of schools did not report graduation rates.
 - We need to know if we are missing any important type of data for unreported graduation rates.
 - Is the type of schools not reporting graduation rates similar to those that do report?

Questions

Thank you for your time!

Sources

- <https://www.businessinsider.com/united-states-regions-new-england-midwest-south-2018-4#-east-south-central-stretches-north-from-mississippi-and-alabama--9>
- <https://educationdata.urban.org/documentation/index.html>