

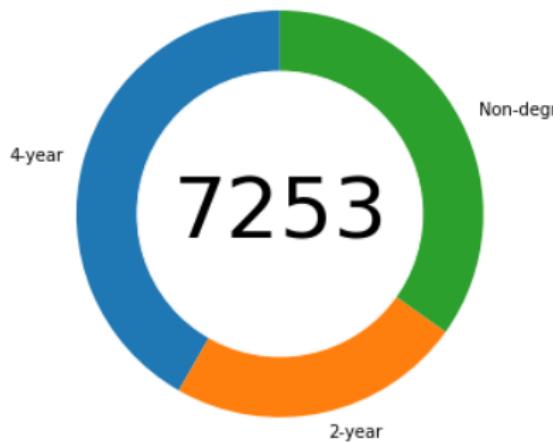
Predicting school closings from the DOE College Scorecard

Kirsten Regier
October 5, 2020



Number of US Postsecondary Institutions

2012-2013



2016-2017

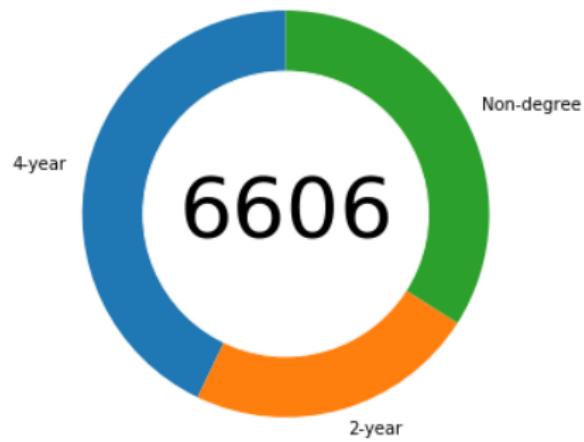


Figure: Distribution of US colleges by degree type for AY 2012-13 (left) and 2016-2017 (right).

Source: U.S. Department of Education, National Center for Education Statistics. (2019). Digest of Education Statistics, 2018 (NCES 2020-009), Chapter 2., <https://nces.ed.gov/fastfacts/display.asp?id=84>

Closing of US postsecondary institutions



postsecondary
institutions closed
between 2012-2013 and
2016-2017.

ITT Technical Institute closed 130+ campuses in 2016, affecting more than 40,000 students and 8,000 faculty

<https://www.insidehighered.com/news/2016/09/07/itt-tech-shuts-down-all-campuses>

Cincinnati Christian University stopped "offering accredited degree programs" in Fall 2019 <https://ccuniversity.edu/a-letter-to-our-students/>

Can we predict which schools are likely to close?

Why does it matter?

- ▶ Trends predict fewer college-going high-school graduates
- ▶ Rising tuition and student loan debt
- ▶ Financial strain from COVID-19

Potential stakeholders:

- ▶ Prospective students (and parents)
- ▶ Faculty, staff, and support services
- ▶ College administrators, boards of trustees, and advisors
- ▶ Accrediting agencies, funding agencies and donors
- ▶ Employers looking to hire graduates

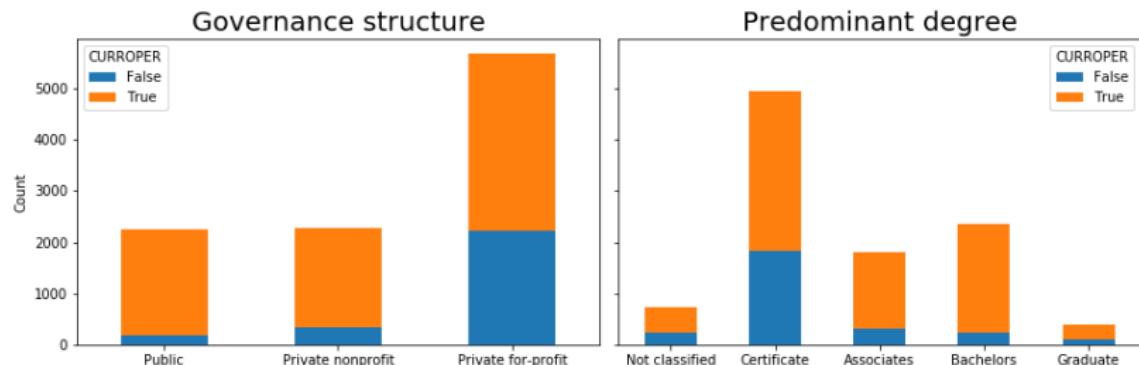
The Data

US Department of Education College Scorecard

(<https://collegescorecard.ed.gov/data/>)

- ▶ **Target variable** - currently operating (CURROPER)
- ▶ 10,200 observations x 32 features
 - ▶ 7441 schools operating in 2013
 - ▶ 2759 closed schools from 2010-2013
- ▶ School identifiers - OPEID, name, location
- ▶ Institutional features - governance structure, predominant degree, highest degree
- ▶ Student features - enrollment, completion rates, median debt
- ▶ Financial info - cost, tuition revenue, institutional expenditure, average faculty monthly salary

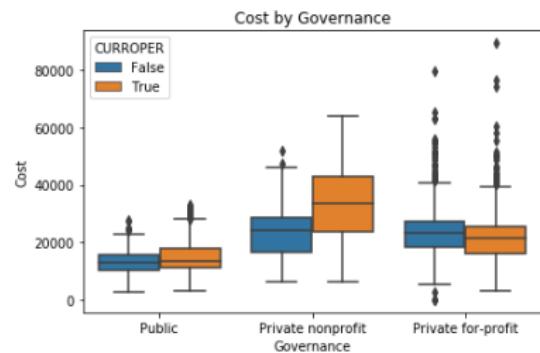
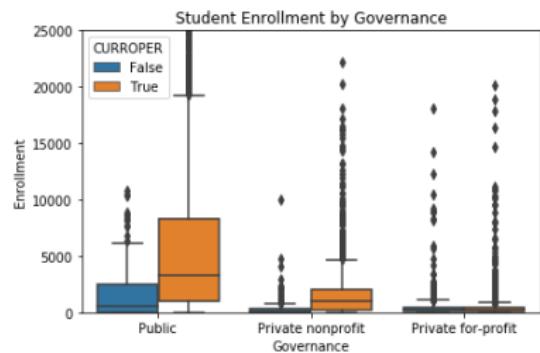
Features of analyzed schools



Distribution of schools in the data by governance (left), predominant degree (right), and whether they are currently operating (CURROPER).

Features of analyzed schools

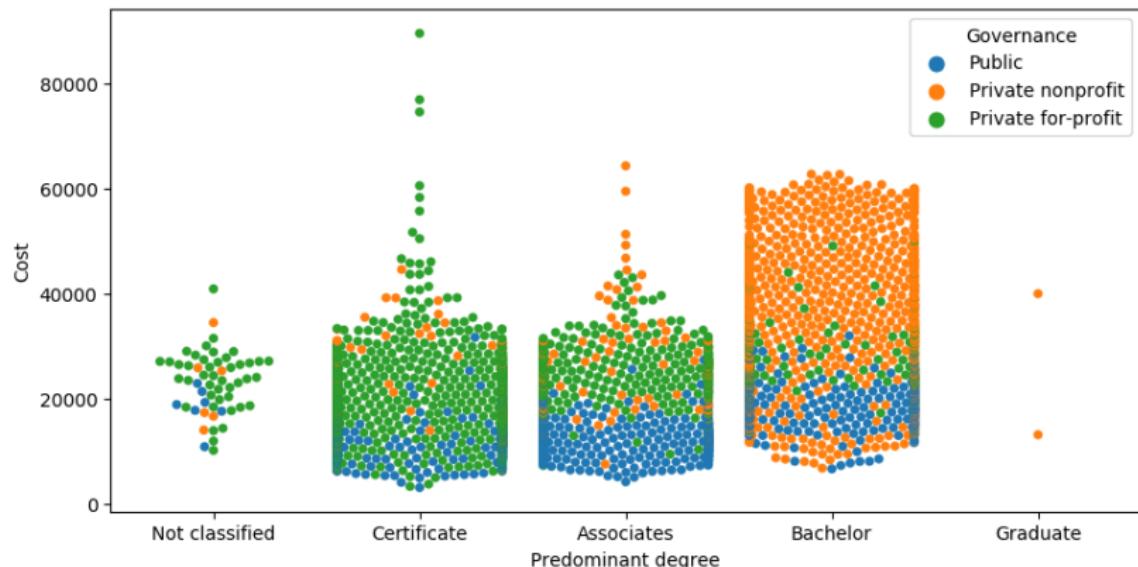
Public schools are larger and less expensive than private schools. Private non-profit schools are slightly larger and more expensive than private for-profit schools, though private for-profit schools show more variation in enrollment and cost than private nonprofit schools.



Enrollment (left) and cost (right) of schools by governance structure and whether they are currently operating (CURROPER).

Features of analyzed schools

Private non-profit schools offer predominantly bachelor's degrees, and are more expensive. Public schools are less expensive than private for-profit schools for all degree types.



Cost by degree type and governance structure.

Recommendations

Use model to predict status of member institutions.

- ▶ Run 180+ member schools through model.
- ▶ Pay attention to schools that are predicted to be 'closed'.

Use model to predict probabilities of school closings.

- ▶ Model generates probability that each school is open.
- ▶ Current model uses 50% as threshold.
- ▶ Can adjust threshold for classification, or examine the probabilities.

Data Processing

Impute missing values using group median

- ▶ Institution name and federal ID
- ▶ Governance structure
- ▶ Predominant degree

Filter data by year

- ▶ Closed schools from 2010-2013
- ▶ Currently operating schools from 2013

Double the number of minority class observations via **resampling** with replacement.

12959 observations x 12 features

Modeling

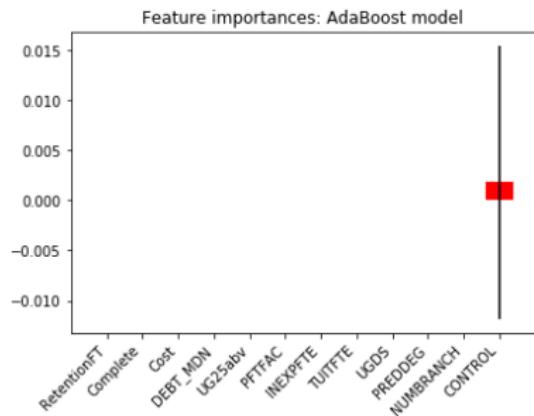
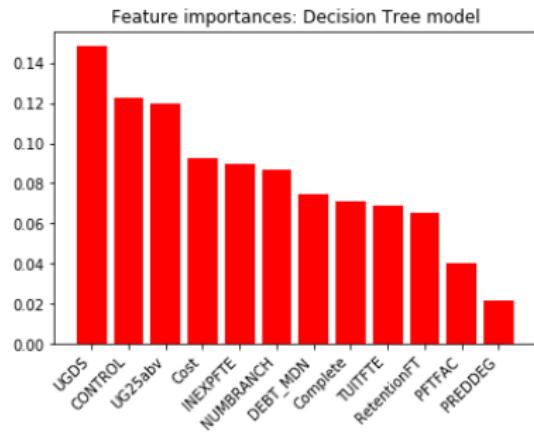
1. Decision Tree model, which feeds
2. AdaBoost model

AdaBoost Model Performance Metrics

		Currently operating		Predicted No	Predicted Yes	Recall
		Actual No	Actual Yes			
Actual	No	1227	30	0.98		
	Yes	69	2163		.97	
Precision		.95		.99		
Accuracy						.97

Table: Confusion matrix and evaluation metrics for the final AdaBoost model with duplicated/resampled observations removed.

Model results



Feature importance levels based on the Decision Tree model (left) and the AdaBoost model (right).

Model results

	Decision Tree Factor	Coefficient	AdaBoost Factor	Coefficient
1	Undergrad enrollment	0.148392	Governance structure	0.001795
2	Governance structure	0.122221	*All other features	null
3	% students over age 25	0.119403		
4	Cost	0.092577		
5	Instructional expenditure	0.089737		

Table: Top 5 predictive features and coefficients

Nonprofit schools - Bachelor's degrees

False negative results

School	Predict	Probab	2013	2020
St. Andrews University	F	0.258114	T	T
Central Methodist University - College of Graduate ...	F	0.252362	T	T
The Robert B Miller College	F	0.440132	T	F

Nonprofit schools

Predicted probability between 50-60%

School	Predict	Probab	2013	2020
Rabbinical Seminary M'kor Chaim	T	0.582	F	No website
Remington College - Dallas	T	0.585	F	T ?
Florida Christian College	T	0.535	F	No website
ICC Technical Institute	T	0.517	F	No website
Remington College of Nursing - Orlando	T	0.581	T	Rename/merge
Alegent Creighton Health School of Radiologic...	T	0.515	T	Rename/merge
CET-Salinas	T	0.531	T	T
National University of Health Sciences	T	0.506	T	T
Baptist Missionary Association Theological Sem...	T	0.582	T	T
Universidad Teologica del Caribe	T	0.544	T	T
Annenberg School of Nursing	T	0.520	T	T

Conclusions

- ▶ Private for-profit schools make up > 50% of the schools in the data. While private for profit schools may be slightly more likely to be predicted as open, there are more closed for-profit schools in the database than other governance types.
- ▶ Schools with more students may be more likely to stay open.

Future directions: Update the model with more recent data, other financial measures (e.g., endowment size, institutional debt, unfunded discount rate), and other data sources (e.g. US News & World Report rankings).