Capstone Final Report

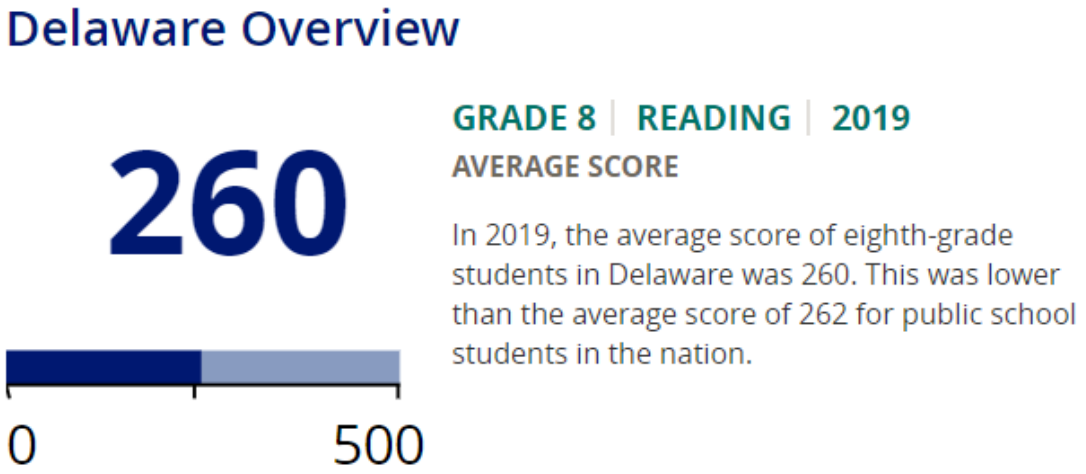
Andrew McLaughlin

Introduction

My capstone project is about staffing levels at schools in the State of Delaware compared to educational outcomes. It was my suspicion that having more staff of various kinds at Delaware schools would improve these outcomes. It is well known that in the state there is a shortage of teachers, but I was interested in staff of all kinds at schools, so they are all included in a model.

Data

My goal was to create a model to project performance at schools in the state of Delaware, and to make a presentation as if I was making it to the higher up decision makers in the Department of Education of the state. Dr. Mark Holodick is the state secretary of education so this model is made specifically for him. I note that on the national education report card, found here: https://www.nationsreportcard.gov/,



Delaware consistently places below the national average in both English and Mathematics. My goal is to improve these scores. From data found on the state of Delaware website on these pages:

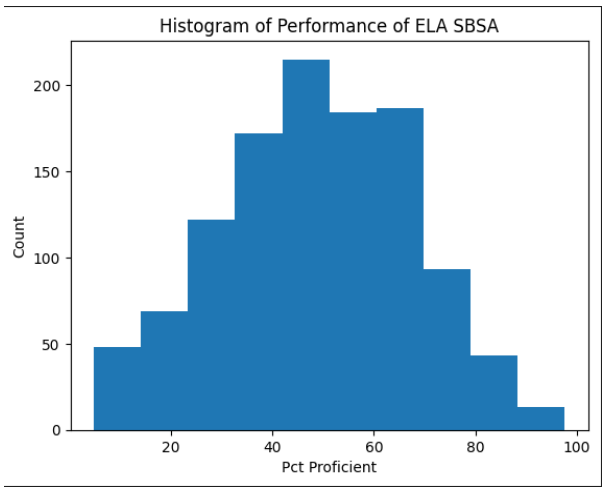
https://data.delaware.gov/Education/Student-Assessment-Performance/ms6b-mt82

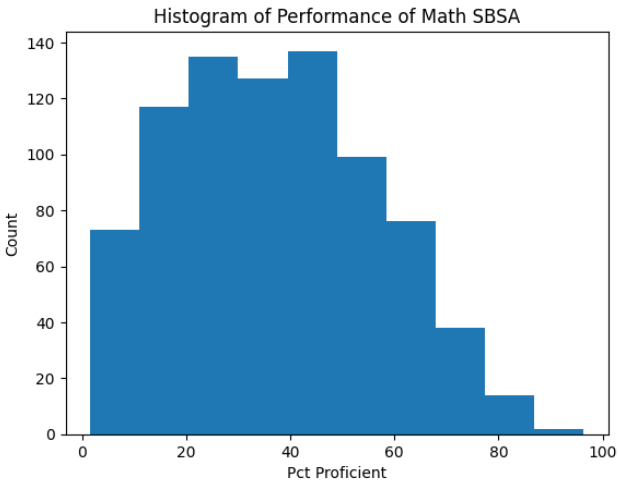
https://data.delaware.gov/Education/Student-Enrollment/6i7v-xnmf

https://data.delaware.gov/Education/Educator-Average-Salary/rv4m-vy79.

Exploratory Data Analysis

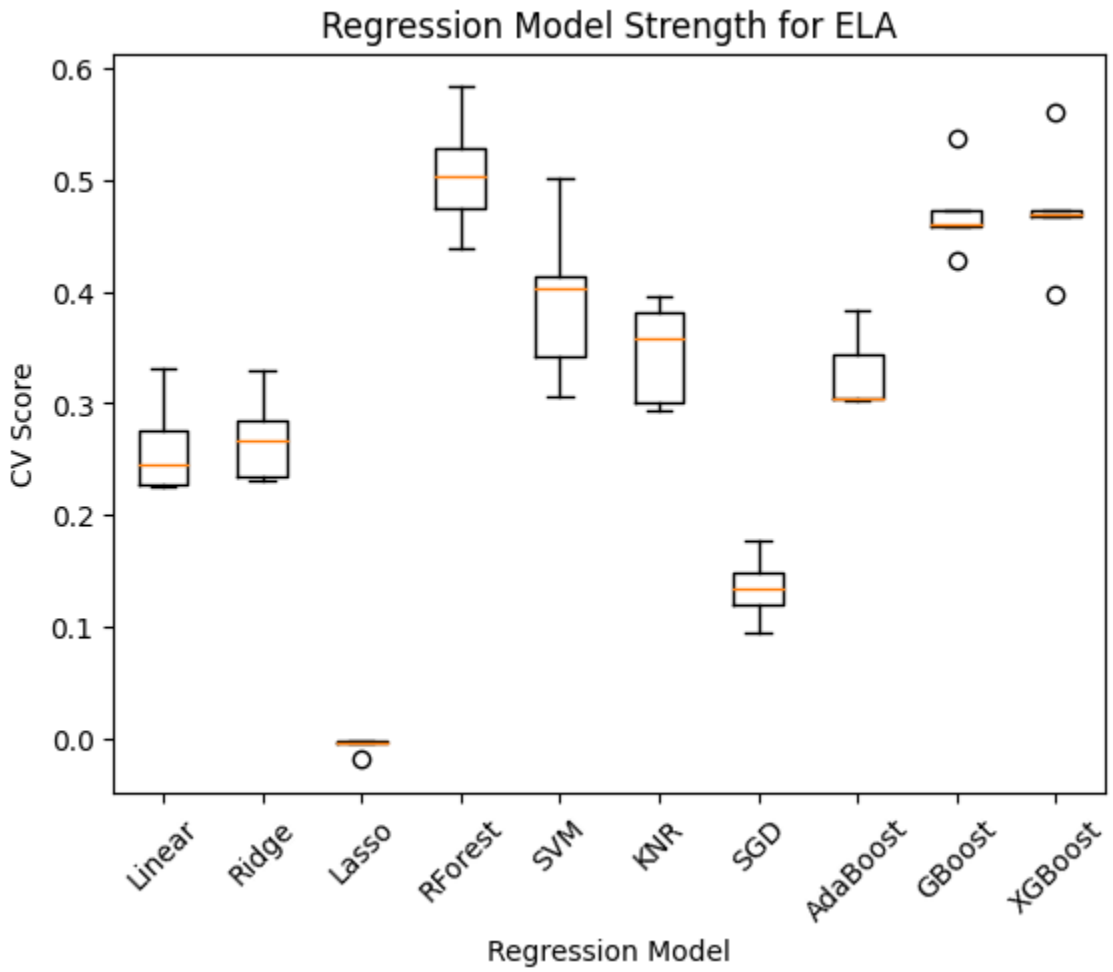
I was able to create a table of data that included student performance and staff/student ratios. In early analysis I saw that the different assessments used throughout the state are not on the same scale so I decided to focus solely on the Smarter Balanced Summative Assessments because they are the most common across grade levels in the state. I also made the decision to split English and Math performance into two different models, because those subject areas have different means and distributions, where English is normally distributed and Math is skewed right.





Model Selection

Now that I had my data I created the model. A quick survey of different model types showed that Random Forest Regression and xgboost regression were showing the best preliminary results in both subject areas.



I decided to use both for both scenarios. After some parameter and hyperparameter tuning, I found that both models worked about equally well, with an R-Square score of about 0.53 on both math models, and about 0.64 on the English models. Ultimately I decided to go with the xgboost models because with its feature importances I believed it

would be easier to make recommendations with it. With that I was ready to apply my models to the statewide data.

Math Recommendation

The xgboost model for math ended up performing poorly when predicting the outcome of statewide data. There could be a few reasons for this. The data making my model may come from slightly older students than the statewide average. This is important because math performance on these tests declines with age. Or it could be that 2022 was a fluke bad year on the test, and without intervention results would improve. Most likely, I think, however, is that COVID learning loss was especially acute in mathematics. Due to the poor performance of the model I decided not to make any suggestions on how to improve math scores. It may be that just staying put is the best option.

ELA Recommendation

With regards to English scores however, I feel much more confident making suggestions. The model predicted state performance much better. One suggestion I can make with the model as to how to improve test scores is to hire 77 more assistant principals in the state and to spread them around the schools evenly. According to the model this would increase proficiency scores from around 40% to around 50%. This makes sense to me as, in my experience, assistant principals help manage the school, and with more of them their time will be divided up between them better, they will have more time to focus on their individual responsibilities and better decisions will be made. Thus the teachers will have better support. One thing that surprised me about the model was that it did not recommend better teacher student ratios. I think this is because of a

few outlier schools in the state. Certain schools have very low student/teacher ratios and bad scores. Like the schools for troubled youth. This may be confusing the model, so I could try in the future to remove schools like this.

Future Models and Conclusion

In the future, I would like to add to my model the different curriculum used, like Savvas or Illustrative for example. This could be implemented in a dummy variable and must improve the model in my mind. Also it is one thing to recommend staff increases but implementing that is another challenge. I know that there is a staff sentiment survey completed every year. If I had the results of that survey, along with churn data, I could make recommendations about how to keep staff. These ideas could be implemented in the future if I had access to that data.

In summary, in order to improve performance on the Smarter Balanced Summative Assessment English tests, hire 77 new assistant principals throughout the state, and that will improve performance, over time, from 40% to 50% proficiency.