Political Science 406: Lab 9: Machine Learning

Due on June 7, 2024

Jason Seawright

Problem 1

Machine Learning Causal Inference

Load King and Zeng's massive dataset on possible predictors of state collapse:

```
kingzeng <- read.csv("kingzeng.csv")
```

The website for these data is at `https://dataverse.harvard.edu/dataset.xhtml?persistentId=hdl:1902.1/RPQIODIANR`, where you can also find a codebook.

We will use the data to analyze the combinations of causes that best predict homicide rates across countries and over time.

Select one possible cause of homicide rates. (If you don't have other ideas, you can use whether a country is democratic or not as a default cause.) Use one or more machine learning algorithms to estimate the effect of that cause on homicide controlling for as many potential confounders as possible. Explain your analysis (as well as your choice of estimator) and interpret your results.