



Neighborhood Crime Prediction

What makes a “safe” neighborhood?

Philip Mark

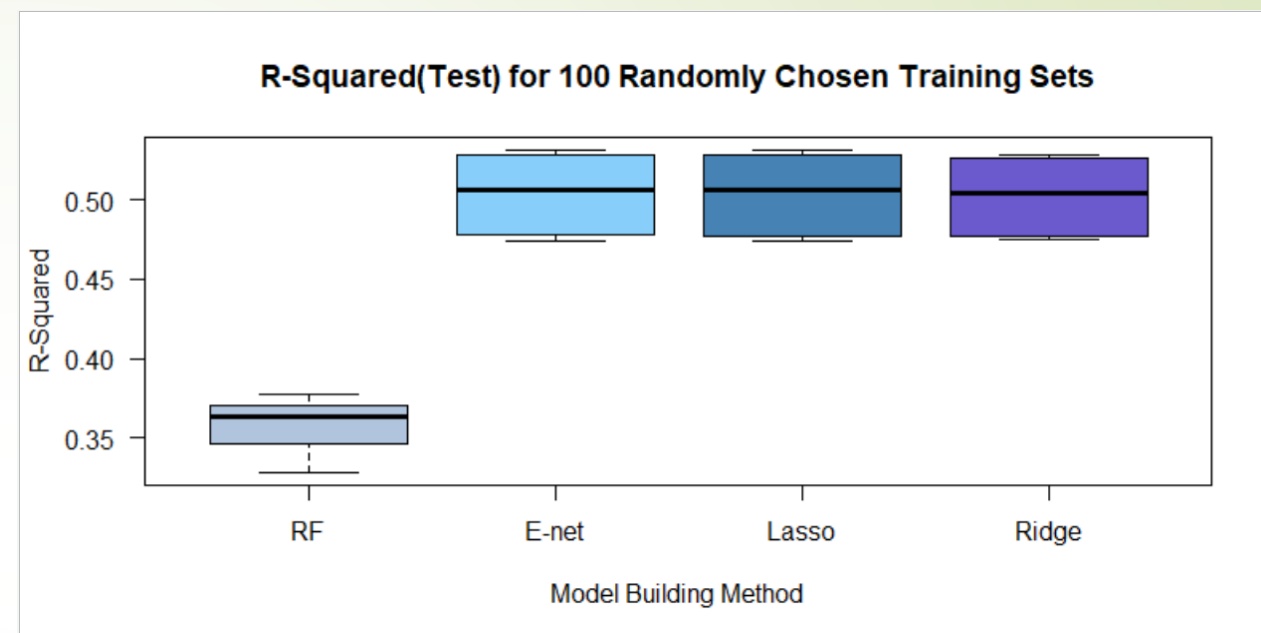
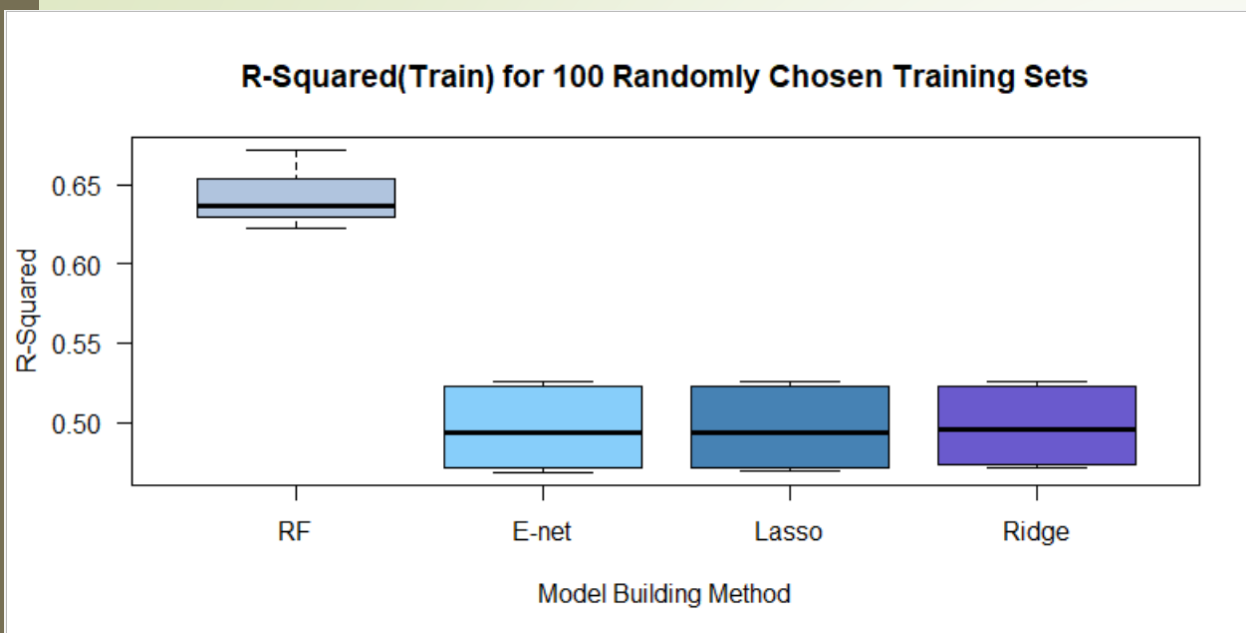
May 2020



The Data

- Crime Rates in US Neighborhoods
- Normalized data capturing 122 different numerical attributes related to a particular community
- 1,994 Communities captured
- Response variable: Violent Crime Rate per Capita

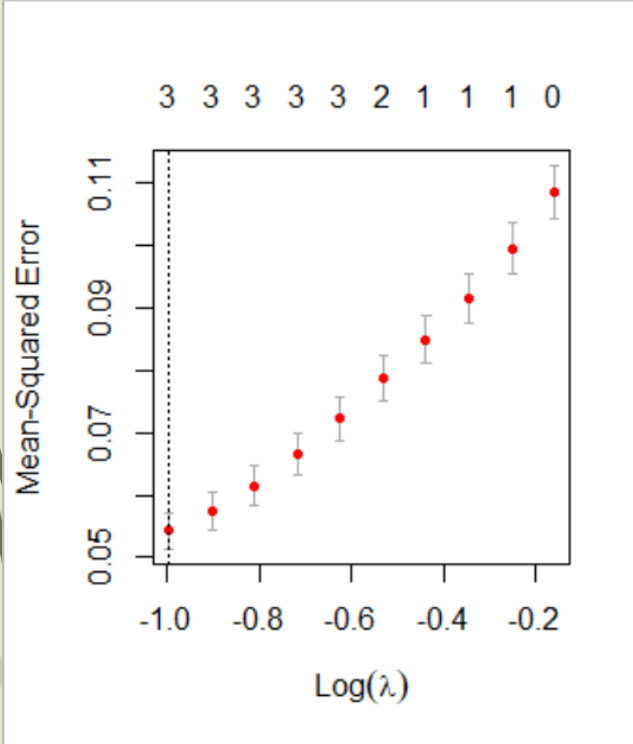
Creator: Michael Redmond (redmond '@' lasalle.edu);
Computer Science; La Salle University; Philadelphia, PA,
19141, USA



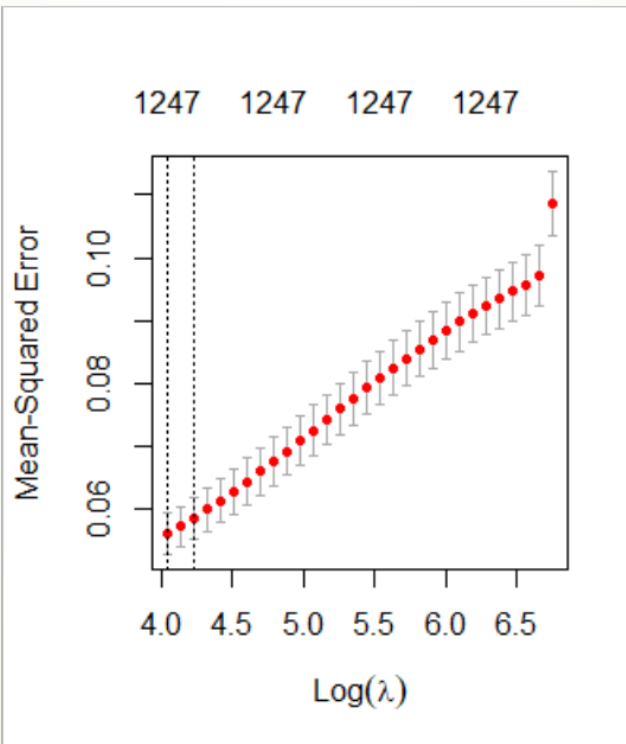
R-Squared Box Plots:
Training Data vs. Test Data

Cross Validation Curves:

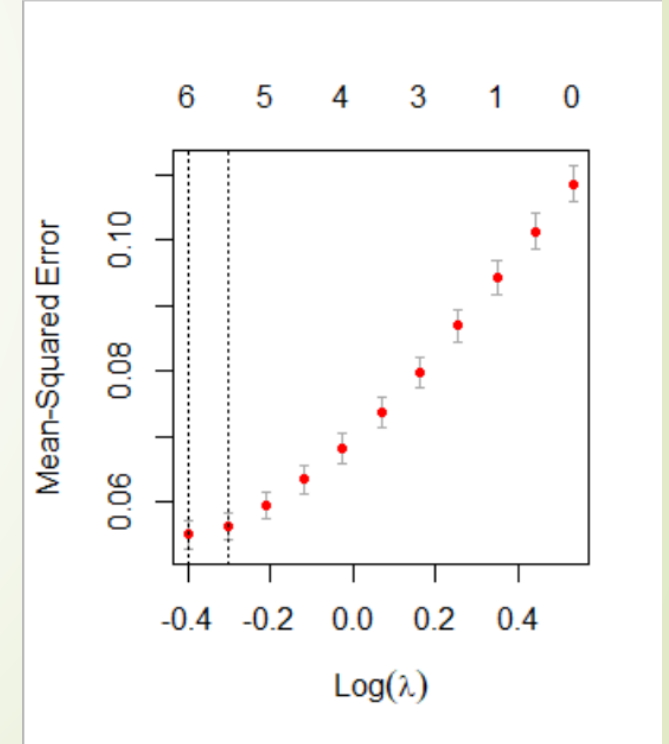
Lasso



Ridge

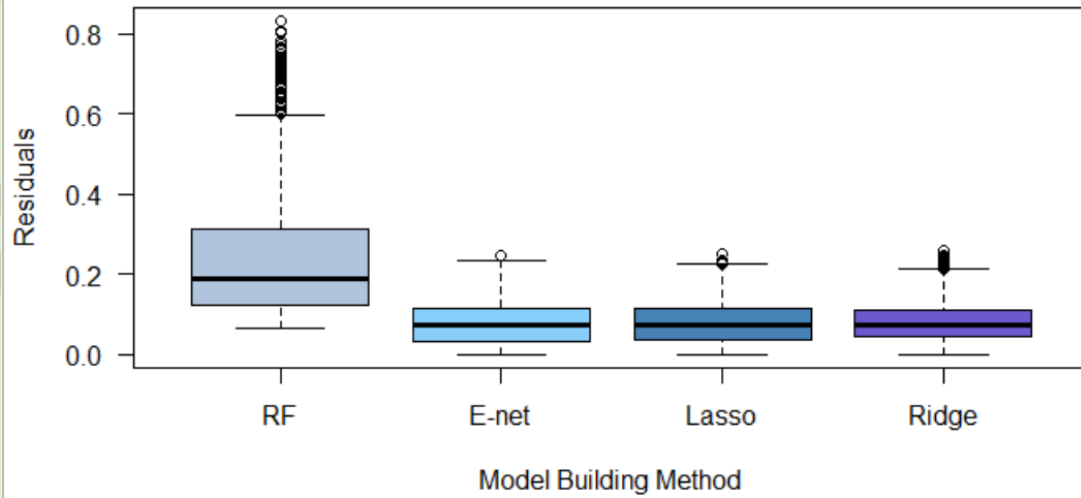


Elastic-net

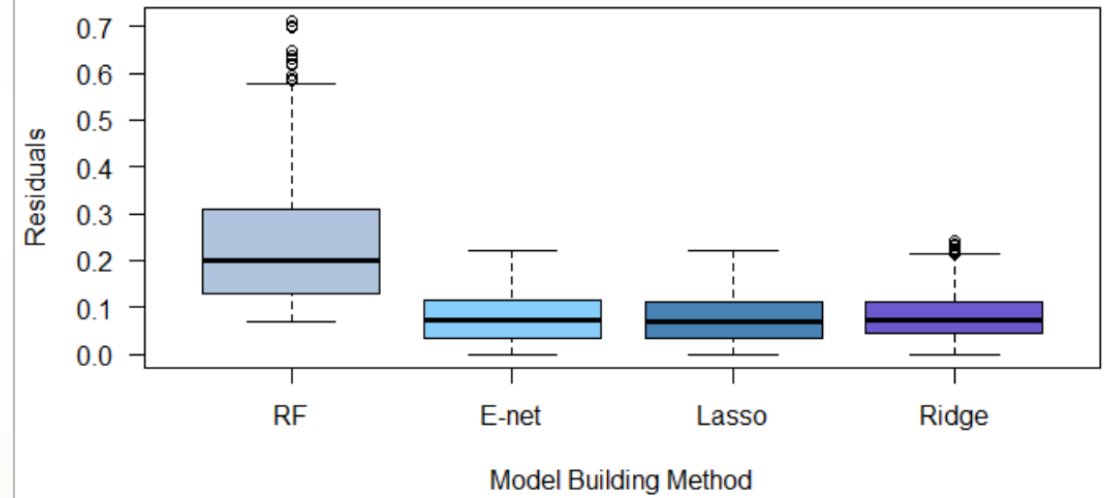


Residuals

Residuals for Training Data

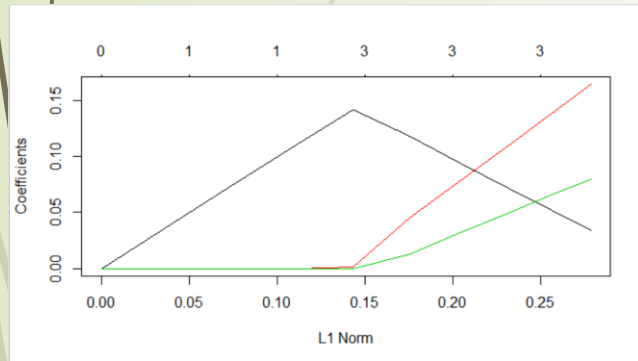


Residuals for Test Data

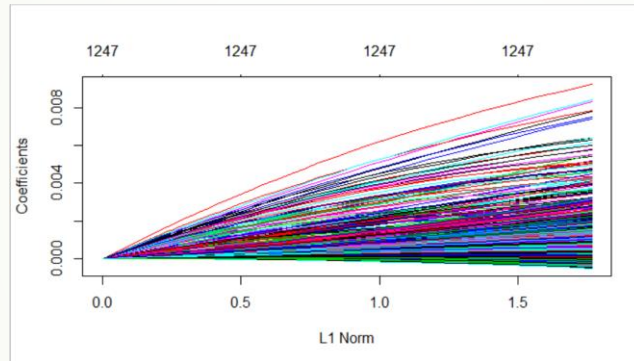


Estimated Coefficients

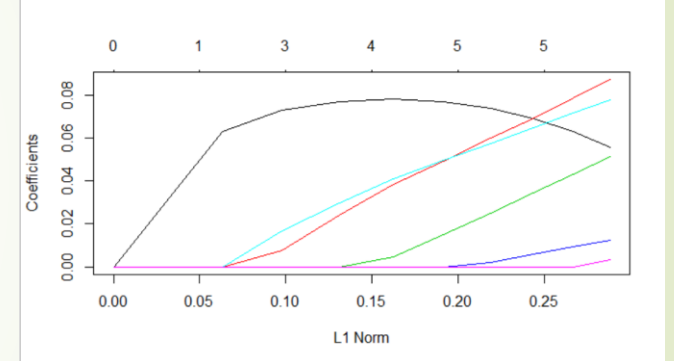
Lasso



Ridge



Elastic-net



Most Significant Variables: 4, 16, 19



Summary

- Models only captured ~50% of the variance in the data
- Indicators of safe neighborhoods:
 - percentage of population that is Caucasian
 - percentage of households with investment / rent income
 - percentage of households with retirement income in 1989
- Conclusion: look for old white neighbors with passive income